

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY

	Type of Search	Vendors and cost where applicable
Searcher: <u>Ed Vearm - Early</u>	NA Sequence (#) <u>2</u>	STN _____
Searcher Phone #: <u>308-4501</u>	AA Sequence (#) _____	Dialog _____
Searcher Location: <u>Biotek lab.</u>	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>9/18/01</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>9/19/01</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep & Review Time: _____	Fulltext _____	Sequence Systems <u>AB5501</u>
Clencal Prep Time: <u>4m.</u>	Patent Family _____	WWW/Internet _____
Online Time: <u>2m.</u>	Other _____	Other (specify) _____

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 16:28:43 ; Search time 214.84 Seconds
(without alignments)
5041.567 Million cell updates/sec

Title: US-09-600-991-1
Perfect score: 1725
Sequence: 1 atgtgggtgaccacactct.....accaccaccaccactag 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	N_Geneseq_0601.*
1:	/SIDS8/gcgdata/geneseq/geneseq/NA1980.DAT.*
2:	/SIDS8/gcgdata/geneseq/geneseq/NA1981.DAT.*
3:	/SIDS8/gcgdata/geneseq/geneseq/NA1982.DAT.*
4:	/SIDS8/gcgdata/geneseq/geneseq/NA1983.DAT.*
5:	/SIDS8/gcgdata/geneseq/geneseq/NA1984.DAT.*
6:	/SIDS8/gcgdata/geneseq/geneseq/NA1985.DAT.*
7:	/SIDS8/gcgdata/geneseq/geneseq/NA1986.DAT.*
8:	/SIDS8/gcgdata/geneseq/geneseq/NA1987.DAT.*
9:	/SIDS8/gcgdata/geneseq/geneseq/NA1988.DAT.*
10:	/SIDS8/gcgdata/geneseq/geneseq/NA1989.DAT.*
11:	/SIDS8/gcgdata/geneseq/geneseq/NA1990.DAT.*
12:	/SIDS8/gcgdata/geneseq/geneseq/NA1991.DAT.*
13:	/SIDS8/gcgdata/geneseq/geneseq/NA1992.DAT.*
14:	/SIDS8/gcgdata/geneseq/geneseq/NA1993.DAT.*
15:	/SIDS8/gcgdata/geneseq/geneseq/NA1994.DAT.*
16:	/SIDS8/gcgdata/geneseq/geneseq/NA1995.DAT.*
17:	/SIDS8/gcgdata/geneseq/geneseq/NA1996.DAT.*
18:	/SIDS8/gcgdata/geneseq/geneseq/NA1997.DAT.*
19:	/SIDS8/gcgdata/geneseq/geneseq/NA1998.DAT.*
20:	/SIDS8/gcgdata/geneseq/geneseq/NA1999.DAT.*
21:	/SIDS8/gcgdata/geneseq/geneseq/NA2000.DAT.*
22:	/SIDS8/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1725	100.0	1725	20	AAX87662 Hepatocyte growth
2	996.6	57.8	1692	20	AAX87663 HGF-MSP hybrid pro
3	857.8	49.7	2190	20	AAX87676 HGF-MSP hybrid pro
4	857.8	49.7	2190	20	AAX87677 HGF-MSP hybrid pro
5	857.8	49.7	2199	12	AAQ15177 Human leukocyte-de
6	857.8	49.7	2199	19	AAV53627 Human leukocyte-de
7	857.8	49.7	2289	11	AAQ06088 Tumour cytotoxic f
8	856.2	49.6	2172	19	AAV61952 Nucleotide sequenc
-9	856.2	49.6	2173	14	AAQ37308 Encodes haematopoi
10	856.2	49.6	2289	13	AAQ21066 TCF II-encoding se
11	851.6	49.4	2172	12	AAQ14182 Plasminogen-like g

12	832.8	48.3	2187	14	AAQ47832 Competative inhibi
13	832.8	48.3	2187	14	AAQ47833 Competative inhibi
14	832.8	48.3	2187	14	AAQ45702 Vascular endotheli
15	832.8	48.3	2214	12	AAQ15176 Human leukocyte-de
16	832.8	48.3	2214	19	AAV53626 Human leukocyte-de
17	831.2	48.2	2184	13	AAQ20049 Human hepatocyte g
18	831.2	48.2	2184	14	AAQ46040 Hepatocyte growth
19	831.2	48.2	2187	12	AAQ10489 Hepatic parenchyma
20	829.6	48.1	2393	12	AAQ14038 Human hepatocyte g
21	828	48.0	2184	15	AAQ56157 Hepatocyte growth
22	825	47.8	873	13	AAQ23662 Truncated hepatocy
23	825	47.8	873	13	AAQ34613 cDNA encoding 34kD
24	825	47.8	873	18	AAT78409 34 kilodalton hepa
25	821.6	47.6	2184	13	AAQ26052 Human HGF. Homo s
26	820	47.5	1661	12	AAQ12398 Human hepatocyte g
27	820	47.5	1661	13	AAQ22144 Human HGF gene par
28	820	47.5	2187	12	AAQ12399 Human hepatocyte g
29	820	47.5	2187	13	AAQ22146 Complete human HGF
30	820	47.5	2187	19	AAV20546 Human recombinant
31	820	47.5	2187	20	AAV15186 Nucleic acid encod
32	682.4	39.6	1400	13	AAQ20745 Encodes alpha-chai
33	680.8	39.5	2187	13	AAQ20255 Encodes alpha- and
34	585	33.9	648	18	AAT78410 Hepatocyte growth
35	302.4	17.5	2394	13	AAQ22143 Human HGF gene par
36	302.4	17.5	2395	12	AAQ12397 Human hepatocyte g
37	170.6	9.9	366	13	AAQ26714 Encodes exon II of
38	159	9.2	2021	15	AAQ79725 Human L5/3 tumour
39	159	9.2	2021	18	AAT62438 Human L5/3 partial
40	159	9.2	2213	18	AAT47149 Macrophage stimula
41	159	9.2	2216	18	AAT47145 Macrophage stimula
42	159	9.2	2216	20	AAV72083 Human MSP cDNA. H
43	159	9.2	2219	15	AAQ79723 Human L5/3 tumour
44	159	9.2	2219	18	AAT62436 Human L5/3 partial
45	159	9.2	2262	15	AAQ79729 Full-length human

ALIGNMENTS

RESULT	1
AAX87662	
ID	AAX87662 standard; cDNA; 1725 BP.
XX	
AC	AAX87662;
XX	
DT	26-OCT-1999 (first entry)
XX	
DE	Hepatocyte growth factor hybrid protein Magic F-1 cDNA.
XX	
KW	Magic Factor-1; human; hepatocyte growth factor; HGF; apoptosis;
KW	chemotherapy; toxicity; therapy; ss.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	sig_peptide 1..63
FT	mat_peptide /*tag= a
FT	64..1725
FT	/*tag= b
XX	
PN	WO9938967-A2.
XX	
PD	05-AUG-1999.
XX	
PF	27-JAN-1999; 99WO-EP00478.
XX	
PR	30-JAN-1998; 98IT-MI00179.
XX	
PA	(DOMP-) DOMPE SPA.
XX	
PI	Caselli G, Collesi C, Comoglio P, Michieli P;
XX	

|||||
Db 1561 cccgacaaggcttgatgataattattgcccgaatcccgatggccagccgagggccatgg 1620
QY 1621 tgctatactcttgacctcacaccgctgggagtactgtgcaattaaacatggcgctgac 1680
Db 1621 tgctatactcttgacctcacaccgctgggagtactgtgcaattaaacatggcgctgac 1680
QY 1681 aaagctgacgacgacgacaaacaccaccaccaccaccaccactag 1725
Db 1681 aaagctgacgacgacgacaaacaccaccaccaccaccaccactag 1725

RESULT 2
AAX87663
ID AAX87663 standard; cDNA; 1692 BP.
XX AAX87663;
XX
DT 26-OCT-1999 (first entry)
XX HGF-MSP hybrid protein Metron Factor-1 cDNA.
DE
XX
KW Metron Factor-1; human; hepatocyte growth factor; HGF;
KW macrophage stimulating protein1 MSP; apoptosis; chemotherapy;
KW toxicity; therapy; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX

Key Location/Qualifiers
FH sig_peptide 1..54
FT /*tag= a
FT mat_peptide 55..1692
FT /*tag= b
XX
PN WO9338967-A2.
XX
XX 05-AUG-1999.
XX
XX 27-JAN-1999; 99WO-EP00478.
XX
XX 30-JAN-1998; 98IT-MI00179.
XX
XX (DOMP-) DOMPE SPA.
XX
XX Caselli G, Collesi C, Comoglio P, Medico E, Michieli P;
XX WPI; 1999-494090/41.
XX P-PSDB; AAY06620.
XX
XX Recombinant proteins from recombination of HGF and MSP structural
XX domains, useful for protection against apoptosis induced by
XX chemotherapeutics
XX
XX Example 1; Page 60-61; 63pp; English.

This DNA sequence codes for Metron F-1 factor (see AAY06620), a recombinant protein composed of the macrophage stimulating protein (MSP) alpha chain signal peptide, hairpin loop and kringle domains 1 and 2 joined via a peptide linker to the hepatocyte growth factor (HGF) hairpin loop and kringle domains 1 and 2 plus a polyhistidine tag. The portions of DNA encoding the various regions of Metron F-1 were obtained by PCR amplification and then recombined to obtain the hybrid sequence. Expression vectors, prokaryotic or eukaryotic host cells and a process for preparing -recombinant proteins from HGF and MSP are claimed. Metron F-1 and other recombinant proteins comprising HGF and/or MSP structural domains are used to prevent or treat chemotherapeutic-induced toxicity such as myelotoxicity, hepatotoxicity, nephrotoxicity, mucotoxicity and neurotoxicity (claimed). They protect cells from -death (apoptosis) induced by chemotherapy of, e.g. tumors. In particular, they can be used for expansion of marrow precursors, to increase proliferation of the haematopoietic precursors or to

CC stimulate their entry in the cycle. By modification of the
CC proteolytic site, hybrid factors can be obtained which are activated
CC by proteases of the endoplasmic reticulum (such as furines) during
CC their synthesis. When the proteolytic site is removed, permanently
CC immature forms of the factors can be obtained, having a potential
CC partial agonistic or antagonistic activity. Different functional
CC domains can be combined so as to modulate the biological effects.
XX
SQ Sequence 1692 BP; 442 A; 457 C; 459 G; 334 T; 0 other;

Query Match 57.8%; Score 996.6; DB 20; Length 1692;
Best Local Similarity 83.2%; Pred. No. 4e-249;
Matches 1166; Conservative 0; Mismatches 214; Indels 21; Gaps 2;

QY 346 tttagacctctatgaaaacaaagactacattagaaaactgcatcattggtaaaggacgcagc 405
Db 292 tgtgacctcttccagaagaagactacgtacggacctgcatcatgaacaatggggttggg 351
QY 406 tacaagggaacagtatctatcactaagagtggcatcaaatgtcagccctggagtccatg 465
Db 352 tacgggggcaccatggccacgaccgtgggtggcctgccctgccaggctggagccacaag 411
QY 466 ataccacacgaacacagcta-----tcggggtaaaagacctacagaaaaactactgt 516
Db 412 ttccggaatgatcacaagtacacgcccactctccggaatggcctggaagagaacttctgc 471
QY 517 cgaatactctcgagggggaagaagggggaccctgggtgtttcacaaagcaatccagaggtacgc 576
Db 472 cgtaacctctgatggcgaccccgagggttccttgggtctacacaaacagaccctgctgtgcgc 531
QY 577 tacgaagtctgtgacattctcctcagttctcgaagtgaatgcacgtgcaatggggag 636
Db 532 ttccagagctgcgggcatcaaatctctccgggagcgcggtgtgtctgtgcaatggcgag 591
QY 637 agttatcgaggtctcatggatcatacagaatcaggcaagatttgtcagcgctgggatcat 696
Db 592 gaataccgcgcggtgtagaccgcacggagtgtagcgagtgccagcgctgggatctt 651
QY 697 cagacaccacacgggcacaaattcttgcctgaaagatatcccgacaaaggctttgatgat 756
Db 652 cagcaccgcacacgaccccttcgagcgcggaagtctcctcgaccaaggtctggacgac 711
QY 757 aattattgccgcaatcccgatggccagccgagcccatgggtgctatactcttgaccctcac 816
Db 712 aactattgccggaatcctgacggctccgagcgccatgggtgctacactacggtaccgcag 771
QY 817 accgctgggagtagtactgtgcaattaaaaacatgcgctgacaaaagc-----ttcg 864
Db 772 atcgagcgagagttctgtgacctcccccgctgcgggtccgaggcacagcccgctcgag 831
QY 865 ggcggtggcggttctgtgtggtggtggttcggtccggtggtggttctctagagggacaaaag 924
Db 832 ggcggtggcggttctgtgtggtggtggttcggtggtggtggttctctagagggacaaaag 891
QY 925 aaaagaagaaatacaaatctcatgaattcaaaaaaatcagcaaaagactaccctaatcaaaaata 984
Db 892 aaaagaagaaatacaaatctcatgaattcaaaaaaatcagcaaaagactaccctaatcaaaaata 951
QY 985 gatccagcactgaagataaaaaaaccaaaaagtgaatactgcagaccaatgtgctaataga 1044
Db 952 gatccagcactgaagataaaaaaaccaaaaagtgaatactgcagaccaatgtgctaataga 1011
QY 1045 tgtactaggaataaaggacttccattccacttgcaggctttgtgttttgataaaagcaaga 1104
Db 1012 tgtactaggaataaaggacttccattccacttgcaggctttgtgttttgataaaagcaaga 1071
QY 1105 aaacaatgcctctggttcccttcaatagcatgtcaagtggagtgaaaaaagaatttggc 1164
Db 1072 aaacaatgcctctggttcccttcaatagcatgtcaagtggagtgaaaaaagaatttggc 1131
QY 1165 catgaatttgacctctatgaaaacaaagactacattagaaaactgcatcattggttaaagga 1224

QY 421 tctatcactaagagtggcatcaaatgtcagccctggagttccatgataccacgaacac 480
|||||
Db 421 tctatcactaagagtggcatcaaatgtcagccctggagttccatgataccacgaacac 480

QY 481 agctatcggggttaagacctacaggaaaaactactgtcgaaatcctcgagggggaagagg 540
|||||
Db 481 agctatcggggttaagacctacaggaaaaactactgtcgaaatcctcgagggggaagagg 540

QY 541 ggaccctggtgtttcacaaagcaatccagaggttacgtacgaagtctgtgacattcctcag 600
|||||
Db 541 ggaccctggtgtttcacaaagcaatccagaggttacgtacgaagtctgtgacattcctcag 600

QY 601 tgttcagaagttaagtcatgacctgacctgcaaatggggagagttatcgagggtctcatgatcat 660
|||||
Db 601 tgttcagaagttaagtcatgacctgacctgcaaatggggagagttatcgagggtctcatgatcat 660

QY 661 acagaatcaggcaagatttgtcagcgtgggtatcatcagacacacacccggcacaaattc 720
|||||
Db 661 acagaatcaggcaagatttgtcagcgtgggtatcatcagacacacacccggcacaaattc 720

QY 721 ttgcctgaaagatatcccgacaaggcgtttgatgataaattattgccgcaatcccgatggc 780
|||||
Db 721 ttgcctgaaagatatcccgacaaggcgtttgatgataaattattgccgcaatcccgatggc 780

QY 781 cagccgagggccatggtgtctatactcttgacccctcacaccccgctgggagtactgtgcaatt 840
|||||
Db 781 cagccgagggccatggtgtctatactcttgacccctcacaccccgctgggagtactgtgcaatt 840

QY 841 aaaacatgcgctgacaaaagt 861
|||||
Db 841 aaaacatgcgctgacaaatact 861

RESULT 4

AAX87677
ID AAX87677 standard; cDNA; 2190 BP.

XX AAX87677;
AC
XX
DT 26-OCT-1999 (first entry)
XX
DE HGF-MSP hybrid protein alphabet-RTKR factor cDNA.
XX
DE Alphabet-RTKR; human; hepatocyte growth factor; HGF; MSP;
KW macrophage stimulating protein; apoptosis; chemotherapy; toxicity;
KW therapy; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT sig_peptide 1..63
FT /*tag= a
FT mat_peptide 63..2190
FT /*tag= b
XX
PN WO9938968-A1.
XX
PD 05-AUG-1999.
XX
PF 28-JAN-1999; 99WO-EF00502.
XX
PR 30-JAN-1998; 98IT-MI00180.
XX
PA -(DOMP-) DOMPE SPA.
XX
PI Caselli G, Collesi C, Comoglio P, Medico E, Michieli P;
XX
DR WPI; 1999-494091/41.
DR -P-PSDB; AAY06622.
XX
PT Recombinant proteins from recombination of HGF and MSP structural

PT domains, useful for protecting cells against apoptosis induced by
chemotherapeutics

XX Example 2; Page 56-58; 63pp; English.

CC This DNA sequence codes for alphabet-RTKR factor (see AAY06622), a
CC recombinant protein composed of the hepatocyte growth factor (HGF)
CC alpha chain (i.e. the signal sequence, hairpin loop and kringles
CC 1-4), the natural cleavage site of HGF plus a furine protease
CC cleavage site, the macrophage stimulating protein (MSP) beta chain
CC and a polyhistidine tag sequence. This structure allows the
CC recombinant protein to interact with both the HGF receptor (Met)
CC and the MSP receptor (Ron) and thereby induce biological responses
CC which are synergistic and selective compared with the natural
CC factors and truncated forms of the proteins. The portions of DNA
CC encoding the various portions of alphabet-RTKR are obtained by PCR
CC amplification of HGF or MSP cDNA and then recombined to obtain the
CC hybrid sequence. Expression vectors, prokaryotic or eukaryotic
CC host cells and a process for preparing recombinant proteins from
CC HGF and MSP are claimed. Alphabet-RTKR and other HGF-MSP
CC recombinant proteins are used for the prevention or treatment of
CC chemotherapeutic-induced toxicity such as myelotoxicity,
CC hepatotoxicity, nephrotoxicity, mucotoxicity and neurotoxicity
CC (claimed). They protect cells from programmed cell death
CC (apoptosis) induced by chemotherapy of, e.g. tumors. In particular,
CC they can be used for expansion of marrow precursors, to increase
CC proliferation of the haematopoietic precursors or to stimulate
CC their entry in the cycle. By modification of the proteolytic
CC site, hybrid factors can be obtained which are activated by
CC proteases of the endoplasmic reticulum (such as furines) during
CC their synthesis. When the proteolytic site is removed, permanently
CC immature forms of the factors can be obtained, having a potential
CC partial agonistic or antagonistic activity. Different functional
CC domains can be combined so as to modulate the biological effects.

XX
SQ Sequence 2190 BP; 612 A; 528 C; 544 G; 506 T; 0 other;

Query Match 49.7%; Score 857.8; DB 20; Length 2190;
Best Local Similarity 99.8%; Pred. No. 4.8e-213;
Matches 859; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 atgtgggtgaccaaaactcctgcagccctgctgctgcagcatgtcctcctgcctcctc 60
|||||
Db 1 atgtgggtgaccaaaactcctgcagccctgctgctgcagcatgtcctcctgcctcctc 60

QY 61 ctgctcccatcgccatccctctatgcagaggggacaaaaggaaaagaataacaattcat 120
|||||
Db 61 ctgctcccatcgccatccctctatgcagaggggacaaaaggaaaagaataacaattcat 120

QY 121 gaattcaaaaaaatcagcaaaagactaccctaatcaaaatagatccagcactgaagataaaa 180
|||||
Db 121 gaattcaaaaaaatcagcaaaagactaccctaatcaaaatagatccagcactgaagataaaa 180

QY 181 accaaaaaagtgaatactactgcagaccatgtgctaataatagatgtactaggaataaaggactt 240
|||||
Db 181 accaaaaaagtgaatactactgcagaccatgtgctaataatagatgtactaggaataaaggactt 240

QY 241 ccattcacttgcaggcttttgtttttgataaaagcaaaaaaatgcctctgttcccc 300
|||||
Db 241 ccattcacttgcaggcttttgtttttgataaaagcaaaaaaatgcctctgttcccc 300

QY 301 ttcaatagcatgtcaagtgagtgaaaaagaatttggccatgaatttgaccttatgaa 360
|||||
Db 301 ttcaatagcatgtcaagtgagtgaaaaagaatttggccatgaatttgaccttatgaa 360

QY 361 aacaaagactacattagaaaactgcatcattgtgtaaaaggacgcagctacaagggaacagta 420
|||||
Db 361 aacaaagactacattagaaaactgcatcattgtgtaaaaggacgcagctacaagggaacagta 420

QY 421 tctatcactaagagtggcatcaaatgtcagccctggagttccatgataccacgaacac 480
|||||
Db 421 tctatcactaagagtggcatcaaatgtcagccctggagttccatgataccacgaacac 480

Db 808 cagccgaggccatgtgtctatactcttgaccctcacaccccgctgggagtactgtgcaatt 867

QY 841 aaaacatgcgctgacaaagct 861
|||||

Db 868 aaaacatgcgctgacaaact 888

RESULT 6

AAV53627

ID AAV53627 standard; cDNA; 2199 BP.

XX AAV53627;

DT 21-DEC-1998 (first entry)

XX Human leukocyte-derived hepatocyte growth factor cDNA clone HLC2.

DE Human leukocyte-derived hepatocyte growth factor; HGF; human; liver; hepatoma;

XX Hepatocyte-derived growth factor; HGF; human; liver; hepatoma;

KW diagnosis; therapy; ss.

KW Homo sapiens.

OS

XX

FH Key Location/Qualifiers

FT CDS 28..2199

FT /*tag= a

XX

PN EP859009-A2.

XX

PD 19-AUG-1998.

XX

PF 07-JUN-1991; 91EP-0109369.

XX

PR 11-JUN-1990; 90JP-0152474.

XX

PA (NAKA/) NAKAMURA T.

XX

PI Asami O, Hagiya M, Ihara I, Nakamura T, Sakaguchi M;

PI Seki T, Shimizu S, Shimonishi M;

XX

DR WPI; 1998-429650/37.

DR P-PSDB; AAW59923.

XX

PT Preparation of recombinant hepatocyte growth factor polypeptide - by

PT culturing mammalian cells transformed with vector containing human

PT leukocyte-derived HGF gene

XX

PS Example 1; Fig 3a-c; 30pp; English.

XX

CC cDNA clone HCL2 codes for a human leukocyte-derived hepatocyte

CC growth factor (HGF, see AAW59923). It was isolated from a human

CC leukocyte cDNA library using a fragment of human liver-derived HGF

CC cDNA as probe. Another clone, HLC3 (see AAV53626), was also

CC isolated. HCL2 differs from HCL3 by having nucleotides 510-524 of

CC HCL3 deleted. A claimed method of producing a HGF comprises:

CC transforming mammalian cells (preferably CHO cells) with a

CC recombinant expression vector that has a promoter (preferably the

CC SV40 early promoter), a gene coding for human leukocyte-derived

CC HGF, and the dihydrofolate reductase gene; (b) culturing the

CC transformed cells in the presence of successively elevated

CC concentrations of methotrexate; and (c) recovering the HGF

CC polypeptide from the culture supernatant. HGF polypeptides may be

CC used as hepatocyte cultivation reagents, liver regeneration

CC promoters, in basic research on liver function, research on the

CC action of various hormones and drugs on hepatocytes, research on

CC the carcinogenesis mechanism of hepatoma, clinical diagnostic

CC -reagents using an antibody against the polypeptide and therapeutic

CC drugs for liver disease.

XX

SQ Sequence 2199 BP; 678 A; 466 C; 503 G; 552 T; 0 other;

Query Match 49.7%; Score 857.8; DB 19; Length 2199;
Best Local Similarity 99.8%; Pred. No. 4.8e-213;

Matches 859; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 atgtgggtgaccaaactcctgccagccctgctgctgcagcatgtcctcctgcattcctc 60
|||||

Db 28 atgtgggtgaccaaactcctgccagccctgctgctgcagcatgtcctcctgcattcctc 87
|||||

QY 61 ctgctccccatcgccatccccctatgcagaggggacaaaaggaaaagatacaattcat 120
|||||

Db 88 ctgctccccatcgccatccccctatgcagaggggacaaaaggaaaagatacaattcat 147
|||||

QY 121 gaattcaaaaaaatcagcaaaagactaccctaatcaaaaatagatccagcactgaagataaaa 180
|||||

Db 148 gaattcaaaaaaatcagcaaaagactaccctaatcaaaaatagatccagcactgaagataaaa 207
|||||

QY 181 accaaaaagtgaatactgcagaccacaaatgtgctaataagatgtactagataaaggactt 240
|||||

Db 208 accaaaaagtgaatactgcagaccacaaatgtgctaataagatgtactagataaaggactt 267
|||||

QY 241 ccattcacttgcaaggcttttgttttttgataaaaagcaagaaaacaatgcctctgttcccc 300
|||||

Db 268 ccattcacttgcaaggcttttgttttttgataaaaagcaagaaaacaatgcctctgttcccc 327
|||||

QY 301 ttcaatagcatgtcaagtggagtgaaaaaagaatttggcccatgaatttgacctctatgaa 360
|||||

Db 328 ttcaatagcatgtcaagtggagtgaaaaaagaatttggcccatgaatttgacctctatgaa 387
|||||

QY 361 aacaaagactacattagaaaactgcatacttgtaaaaggacgcagctacaagggaacagta 420
|||||

Db 388 aacaaagactacattagaaaactgcatacttgtaaaaggacgcagctacaagggaacagta 447
|||||

QY 421 tctatcactaagagtggtgcatcaaatgtcagccctggagttccatgataccacacgaacac 480
|||||

Db 448 tctatcactaagagtggtgcatcaaatgtcagccctggagttccatgataccacacgaacac 507
|||||

QY 481 agctatcggggttaaagacctacaggaaaactactgtcgaaatcctcgaggggagaagg 540
|||||

Db 508 agctatcggggttaaagacctacaggaaaactactgtcgaaatcctcgaggggagaagg 567
|||||

QY 541 ggacctgtgtttcacaaagcaatccagagggtacgtacgaagtctgtacattcctcag 600
|||||

Db 568 ggacctgtgtttcacaaagcaatccagagggtacgtacgaagtctgtacattcctcag 627
|||||

QY 601 tgttcagaagttgaatgcatagcctgcaatggggagagttatcgagggtcctgcattcat 660
|||||

Db 628 tgttcagaagttgaatgcatagcctgcaatggggagagttatcgagggtcctgcattcat 687
|||||

QY 661 acagaatcaggcaagatttgtcagcgtgggatcatcagacaccacacccgcacaaattc 720
|||||

Db 688 acagaatcaggcaagatttgtcagcgtgggatcatcagacaccacacccgcacaaattc 747
|||||

QY 721 ttgcctgaaagatatcccgacaaaggcctttgatgataaattattgccgcaatcccgatggc 780
|||||

Db 748 ttgcctgaaagatatcccgacaaaggcctttgatgataaattattgccgcaatcccgatggc 807
|||||

QY 781 cagccgaggccatggtgctatactcttgaccctcacacccgctgggagtactgtgcaatt 840
|||||

Db 808 cagccgaggccatggtgctatactcttgaccctcacacccgctgggagtactgtgcaatt 867
|||||

QY 841 aaaacatgcgctgacaaagct 861
|||||

Db 868 aaaacatgcgctgacaaact 888
|||||

RESULT 7

AAQ06088

ID AAQ06088 standard; cDNA; 2289 BP.

XX

AC AAQ06088;

XX

DT 31-JAN-1991 (first entry)

XX

DE Tumour cytotoxic factor II gene.

XX


```
XX Claim 1; Fig 6A-D; 19pp; English.
PS
XX
CC The nucleotide sequence of human plasminogen-like growth factor
CC PLGF (2172 bp) is used in the stimulation of the growth of endothelial
CC cells or epithelial cells other than hepatocytes. This involves
CC contacting the cells with the PLGF protein or its allelic or splice
CC variants. The growth factor is useful for stimulating wound healing in
CC endothelial or epithelial tissue.
XX
SQ Sequence 2172 BP; 674 A; 450 C; 498 G; 550 T; 0 other;

Query Match          49.6%; Score 856.2; DB 19; Length 2172;
Best Local Similarity 99.7%; Pred. No. 1.2e-212;
Matches 858; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 atgtgggtgaccaaactcctgccagccctgctgtgcagcatgtcctcctgcattcctc 60
Db 1 atgtgggtgaccaaactcctgccagccctgctgtgcagcatgtcctcctgcattcctc 60

QY 61 ctgctccccatgccatccccctatgcagaggggacaaaagaaagataataattcat 120
Db 61 ctgctccccatgccatccccctatgcagaggggacaaaagaaagataataattcat 120

QY 121 gaattcaaaaaatcagaaagactaccctaatacaaaatagatccagcactgaagataaaa 180
Db 121 gaattcaaaaaatcagaaagactaccctaatacaaaatagatccagcactgaagataaaa 180

QY 121 gaattcaaaaaatcagaaagactaccctaatacaaaatagatccagcactgaagataaaa 180
Db 121 gaattcaaaaaatcagaaagactaccctaatacaaaatagatccagcactgaagataaaa 180

QY 181 accaaaaaagtgaataactgcagaccacaaatgtgctaataatgtaggaataaaggactt 240
Db 181 accaaaaaagtgaataactgcagaccacaaatgtgctaataatgtaggaataaaggactt 240

QY 181 accaaaaaagtgaataactgcagaccacaaatgtgctaataatgtaggaataaaggactt 240
Db 181 accaaaaaagtgaataactgcagaccacaaatgtgctaataatgtaggaataaaggactt 240

QY 241 ccattcacttgcaaggctttgttttgataaagcaagaaaaacaatgcctctgttcccc 300
Db 241 ccattcacttgcaaggctttgttttgataaagcaagaaaaacaatgcctctgttcccc 300

QY 301 ttcaatagatgtcaagtgagtgagtgaaaaaagaatttggccatgaatttgaccttatgaa 360
Db 301 ttcaatagatgtcaagtgagtgagtgaaaaaagaatttggccatgaatttgaccttatgaa 360

QY 361 acaaaagactacattagaaaactgcattctgttaaaggacgcagctacaagggaacagta 420
Db 361 acaaaagactacattagaaaactgcattctgttaaaggacgcagctacaagggaacagta 420

QY 421 tctatcactaagagtggcatcaaatgtcagccctggagttccatgataccacacgaacac 480
Db 421 tctatcactaagagtggcatcaaatgtcagccctggagttccatgataccacacgaacac 480

QY 481 agctatcgggggtaaagacctaagaaaactactgtcgaaaatctctcaggggaagaagg 540
Db 481 agctatcgggggtaaagacctaagaaaactactgtcgaaaatctctcaggggaagaagg 540

QY 541 ggaccctggtgtttcacaaagcaatccagaggtacgctacgaaagtgtgtgacattcctcag 600
Db 541 ggaccctggtgtttcacaaagcaatccagaggtacgctacgaaagtgtgtgacattcctcag 600

QY 601 tgttcagaagttgaatgcattgacctgcaatggggagagttatcgaggtctcatggatcat 660
Db 601 tgttcagaagttgaatgcattgacctgcaatggggagagttatcgaggtctcatggatcat 660

QY 661 acagaatcaggcaagatttgtcagcgtgggatcatcagacaccacaccgggcacaaattc 720
Db 661 acagaatcaggcaagatttgtcagcgtgggatcatcagacaccacaccgggcacaaattc 720

QY 721 ttgcctgaaagatatcccgacaaagggcttttgatgataaattattgccgcaatcccgatggc 780
Db 721 ttgcctgaaagatatcccgacaaagggcttttgatgataaattattgccgcaatcccgatggc 780

QY 781 cagccgagggccatggtgctatactcttgacccctcacaccccgctgggagtagtgcaatt 840
Db 781 cagccgagggccatggtgctatactcttgacccctcacaccccgctgggagtagtgcaatt 840
```

```
QY 841 aaaacatgcgctgacaaaagct 861
Db 841 aaaacatgcgctgacaaatact 861

RESULT 9
AAQ37308
ID AAQ37308 standard; DNA; 2173 BP.
XX
AC AAQ37308;
XX 16-JUN-1993 (first entry)
XX Encodes haematopoietic stem cell multiplier.
XX Bone marrow deficiencies; cancer therapy; tumour; carcinoma;
KW bone marrow transplants.
XX Key Location/Qualifiers
FH CDS 1..2169
FT /*tag= a
XX WO9303061-A.
PN 18-FEB-1993.
XX 24-JUL-1992; 92WO-JP00949.
PR 26-JUL-1991; 91JP-0187470.
PR 26-JUL-1991; 91JP-0187481.
XX (TORA ) TORAY IND INC.
XX Kawano G, Kojima K, Komiyama A, Kubo T, Nakahata T;
PI Sano E, Sudot, Tanaka R;
XX WPI; 1993-076441/09.
DR P-PSDB; AAR32710.
XX Haematopoietic stem cell multiplier comprising IL-3 and IL-7 -
PT used for treatment and prevention of bone marrow disorders e.g.
PT after cancer therapy or bone marrow transplants
XX Disclosure; Page 60; 90pp; Japanese.
XX This sequence encodes a haematopoietic stem cell multiplier of mol.
CC wt. 60,000, and having N-terminal sequence AAR32709.
XX Sequence 2173 BP; 671 A; 452 C; 502 G; 548 T; 0 other;
SQ

Query Match          49.6%; Score 856.2; DB 14; Length 2173;
Best Local Similarity 99.7%; Pred. No. 1.2e-212;
Matches 858; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 atgtgggtgaccaaactcctgccagccctgctgtgcagcatgtcctcctgcattcctc 60
Db 1 atgtgggtgaccaaactcctgccagccctgctgtgcagcatgtcctcctgcattcctc 60

QY 61 ctgctccccatgccatccccctatgcagaggggacaaaagaaagataataattcat 120
Db 61 ctgctccccatgccatccccctatgcagaggggacaaaagaaagataataattcat 120

QY 121 gaattcaaaaaatcagcaaaagactaccctaatacaaaatagatccagcactgaagataaaa 180
Db 121 gaattcaaaaaatcagcaaaagactaccctaatacaaaatagatccagcactgaagataaaa 180

QY 181 accaaaaaagtgaataactgcagaccacaaatgtgctaataatgtaggaataaaggactt 240
Db 181 accaaaaaagtgaataactgcagaccacaaatgtgctaataatgtaggaataaaggactt 240

QY 241 ccattcacttgcaaggctttgttttgataaagcaagaaaaacaatgcctctgttcccc 300
Db 241 ccattcacttgcaaggctttgttttgataaagcaagaaaaacaatgcctctgttcccc 300
```

QY 721 ttgcctgaaagatatcccgacaaagggcttgatgataattattgcccgaatccccgatggc 780
|||||
Db 798 ttgcctgaaagatatcccgacaaagggcttgatgataattattgcccgaatccccgatggc 857
QY 781 cagccgaggccatggtgctatactcttgaccctcacaccccgctgggagtactgtgcaatt 840
|||||
Db 858 cagccgaggccatggtgctatactcttgaccctcacaccccgctgggagtactgtgcaatt 917
QY 841 aaaacatgcctgacaaaagct 861
|||||
Db 918 aaaacatgcctgacaaatact 938

RESULT 11
AAQ14182
ID AAQ14182 standard; DNA; 2172 BP.
XX
AC AAQ14182;
XX
DT 10-JAN-1992 (first entry)
XX
DE Plasminogen-like growth factor gene.
XX
KW Mitogenic activity; PLGF; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT sig_peptide 1..63 /*tag= a
FT mat_peptide 64..2172 /*tag= b
XX
PN US7582063-A.
XX
PD 17-SEP-1991.
XX
PF 14-SEP-1990; 90US-0582063.
XX
PR 14-SEP-1990; 90US-0582063.
XX
PA (USSH) NAT INST OF HEALTH.
XX
PI Rubin JS, Chan AML, Aaronson SA;
XX
DR WPI; 1991-317957/43.
DR P-PSDB; AAR14307.
XX
PT New plasminogen-like growth factor protein - having potent
PT mitogenic activity on melanocytes, epithelial cells and
PT endothelial cells
XX
PS Disclosure; Fig 6; 33pp; English.
XX
CC The DNA sequence is that of a plasminogen-like growth factor (PGLF)
CC gene. The DNA encoding PGLF can be obtd. from a DNA library of M246
CC human embryonic lung fibroblasts. The sequence may be used in the
CC determin. of the level of mRNA transcripts encoding the protein
CC which may be useful in the diagnosis of malignancy.
XX
SQ Sequence 2172 BP; 675 A; 448 C; 499 G; 550 T; 2 other;

Query Match 49.4%; Score 851.6; DB 12; Length 2172;
Best Local Similarity 99.2%; Pred. No. 1.9e-211;
Matches 854; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 atgtgggtgacaaactcctgccagccctgctgctgcagcatgtcctcctgcattcctc 60
|||||
Db 1 atgtgggtgacaaactcctgccagccctgctgctgcagcatgtcctcctgcattcctc 60
QY 61 ctgctcccccatgccatccccctatgcagaggggacaaaggaaaagaataacaattcat 120
|||||

Db 61 ctgctcccccatgccatccccctatgcagaggggacaaaggaaaagaataacaattcat 120
QY 121 gaattcaaaaaatcagcaagactaccctaatcaaaatagatccagcactgaagataaaa 180
|||||
Db 121 gaattcaaaaaatcagcaagactaccctaatcaaaatagatccagcactgaagataaaa 180
QY 181 accaaaaaagtgaatactgcagaccaatgtgctaataatagatggaataaaaggactt 240
|||||
Db 181 accaaaaaagtgaatactgcagaccaatgtgctaataatagatggaataaaaggactt 240
QY 241 ccattcacttgcaaggctttgttttgataaaagcaaaaaaatgcctctggttcccc 300
|||||
Db 241 ccattcacttgcaaggctttgttttgataaaagcaaaaaaatgcctctggttcccc 300
QY 301 ttcaatagcatgtcaagtggagtgaataaaagaatttggccatgaatttgacctctatgaa 360
|||||
Db 301 ttcaatagcatgtcaagtggagtgaataaaagaatttggccatgaatttgacctctatgaa 360
QY 361 acaaaagactacattagaaaactgcattcattggttaaaggcagctacaagggaacagta 420
|||||
Db 361 acaaaagactacattagaaaactgcattcattggttaaaggcagctacaagggaacagtn 420
QY 421 tctatcactaaagagtggcatcaaatgtcagccctggagttccatgataccacacgaacac 480
|||||
Db 421 tctatcactaaagagtggcatcaaatgtcagccctggagttccatgataccacacgaacac 480
QY 481 agctatcgggggttaaagacctacagggaaaactactgtcgaaatmctcgagggggaagagg 540
|||||
Db 481 agctatcgggggttaaagacctacagggaaaactactgtcgaaatmctcgagggggaagagg 540
QY 541 ggaccctggtgtttcacaaagcaatccagaggtacgtacgaaatcctcgagggggaagagg 600
|||||
Db 541 ggaccctggtgtttcacaaagcaatccagaggtacgtacgaaatcctcgacattcctcag 600
QY 601 tgttcagaagttgaatgcacacctgcaatggggagagttatcgaggtctcatggatcat 660
|||||
Db 601 tgttcagaagttgaatgcacacctgcaatggggagagttatcgaggtctcatggatcat 660
QY 661 acagaatcaggcaagatttgcagcgtgggatcatcatcacaccacacccggcacaaattc 720
|||||
Db 661 acagaatcaggcaagatttgcagcgtgggatcatcatcacaccacacagcggcacaaattc 720
QY 721 ttgcctgaaagatatcccgacaaagggctttgatgataaattattgccgcaatccccgatggc 780
|||||
Db 721 ttgcctgaaagatataccgacaaagggtttgatgataaattattgccgcaatccccgatggc 780
QY 781 cagccgaggccatggtgctatactcttgaccctcacaccccgctgggagtactgtgcaatt 840
|||||
Db 781 cagccgaggccatggtgctatactcttgaccctcacaccccgctgggagtactgtgcaatt 840
QY 841 aaaacatgcctgacaaaagct 861
|||||
Db 841 aaaacatgcctgacaaatact 861

RESULT 12
AAQ47832
ID AAQ47832 standard; cDNA to mRNA; 2187 BP.
XX
AC AAQ47832;
XX
DT 14-MAR-1994 (first entry)
XX
DE Competative inhibitor of HGF.
XX
KW HGF; hepatocyte growth factor; transformation; antagonist;
KW liver disease; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..2187
FT /*tag= a

Qy	586	tgtgacattcctcagtggttcagaagttgaatgcacctgcaatggggagaggttatcga	645
Db	628	tgtgacattcctcagtggttcagaagttgaatgcacctgcaatggggagaggttatcga	687
Qy	646	ggtctcatggatcatcacagaatcaggcaagatttgtcagcgcctgggacatcatcacacacca	705
Db	688	ggtctcatggatcatcacagaatcaggcaagatttgtcagcgcctgggacatcatcacacacca	747
Qy	706	caccggcacaaaattcttgctgaaagatatcccgacaaagggtttgatgataaattattgc	765
Db	748	caccggcacaaaattcttgctgaaagatatcccgacaaagggtttgatgataaattattgc	807
Qy	766	cgcaatcccgatggccagccgagggccatgggtgctatactcttgacccctcacaccccgctgg	825
Db	808	cgcaatcccgatggccagccgagggccatgggtgctatactcttgacccctcacaccccgctgg	867
Qy	826	gagtactgtgcaattaaacatgcgctgacaaaagct	861
Db	868	gagtactgtgcaattaaacatgcgctgacaaaagct	903

Search completed: September 18, 2001, 18:33:25
Job time: 7482 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 16:21:47 ; Search time 97.89 Seconds
(without alignments)
3336.007 Million cell updates/sec

Title: US-09-600-991-1
Perfect score: 1725
Sequence: 1 atgtgggtgaccaaactcct.....accaccaccaccaccactag 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	857.8	49.7	2289	1 US-07-838-410-2	Sequence 2, Appli
2	856.2	49.6	2172	4 US-08-030-410-2	Sequence 2, Appli
3	845.8	49.0	2288	1 US-08-290-937B-4	Sequence 4, Appli
4	832.8	48.3	2184	1 US-07-815-333A-1	Sequence 1, Appli
5	159	9.2	2021	1 US-07-882-925A-3	Sequence 3, Appli
6	159	9.2	2021	1 US-08-184-012C-3	Sequence 3, Appli
7	159	9.2	2216	2 US-08-666-082B-2	Sequence 2, Appli
8	159	9.2	2219	1 US-07-882-925A-1	Sequence 1, Appli
9	159	9.2	2219	1 US-08-184-012C-1	Sequence 1, Appli
10	159	9.2	2232	1 US-08-334-177-1	Sequence 1, Appli
11	159	9.2	2232	5 PCT-US95-13830-1	Sequence 1, Appli
12	159	9.2	2262	1 US-07-882-925A-7	Sequence 7, Appli
13	159	9.2	2262	1 US-08-184-012C-7	Sequence 7, Appli
14	157.4	9.1	2219	1 US-08-184-012C-2	Sequence 2, Appli
15	157.4	9.1	2219	1 US-07-882-925A-2	Sequence 2, Appli
16	146	8.5	2188	1 US-07-882-925A-4	Sequence 4, Appli
17	146	8.5	2188	1 US-08-184-012C-4	Sequence 4, Appli
18	136.4	7.9	1284	3 US-08-985-526-24	Sequence 24, Appli
19	136.2	7.9	645	3 US-08-985-526-22	Sequence 22, Appli
20	136	7.9	2497	1 US-08-643-219-12	Sequence 12, Appli
21	136	7.9	2497	2 US-09-131-995-12	Sequence 12, Appli
22	136	7.9	2497	2 US-08-832-087B-12	Sequence 12, Appli
23	136	7.9	2497	3 US-08-851-350-12	Sequence 12, Appli
24	136	7.9	2497	4 US-09-132-154-12	Sequence 12, Appli
25	136	7.9	2679	6 5200340-7	Patent No. 5200340
26	134.4	7.8	2753	1 US-07-854-603-1	Sequence 1, Appli
27	134	7.8	1134	4 US-09-206-059-29	Sequence 29, Appli

28	128.4	7.4	2296	1 US-07-750-080A-18	Sequence 18, Appli
29	128.4	7.4	2296	3 US-08-651-472-18	Sequence 18, Appli
30	55.4	3.2	3398	5 PCT-US95-08493-12	Sequence 12, Appli
31	50.8	2.9	6751	1 US-07-882-925A-5	Sequence 5, Appli
32	50.8	2.9	6751	1 US-08-184-012C-5	Sequence 5, Appli
33	50	2.9	6100	1 US-07-882-925A-6	Sequence 6, Appli
34	50	2.9	6100	1 US-08-184-012C-6	Sequence 6, Appli
c 35	47.8	2.8	7218	1 US-08-232-463-14	Sequence 14, Appli
36	46.8	2.7	175	1 US-08-235-838-3	Sequence 3, Appli
37	46.8	2.7	175	2 US-08-465-473B-3	Sequence 3, Appli
38	46.8	2.7	720	2 US-08-800-198-7	Sequence 7, Appli
39	46.8	2.7	720	3 US-09-296-595-7	Sequence 7, Appli
40	46.8	2.7	748	1 US-08-235-838-4	Sequence 4, Appli
41	46.8	2.7	748	1 US-08-235-838-10	Sequence 10, Appli
42	46.8	2.7	748	2 US-08-465-473B-4	Sequence 4, Appli
43	46.8	2.7	748	2 US-08-465-473B-10	Sequence 10, Appli
44	46.8	2.7	2012	1 US-08-235-838-13	Sequence 13, Appli
45	46.8	2.7	2012	1 US-08-235-838-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1
US-07-838-410-2
; Sequence 2, Application US/07838410
; Patent No. 5328836
; GENERAL INFORMATION:
; APPLICANT: SHIMA, NOBUYUKI
; APPLICANT: HIGASHIO, KANJI
; APPLICANT: NAGAO, MASAYA
; APPLICANT: OOGAKI, FUMIKO
; APPLICANT: TAKAOKA, HIROAKI
; APPLICANT: TSUDA, EISUKE
; TITLE OF INVENTION: PLASMIDS CONTAINING DNA ENCODING THE
; TITLE OF INVENTION: AMINO ACID SEQUENCE OF TCF-II, TRANSFORMED CELLS WITH THE
; TITLE OF INVENTION: PLASMIDS, AND PRODUCTION METHOD OF BIOLOGICALLY ACTIVE
; TITLE OF INVENTION: SUBSTANCE USING THE TRANSFORMED CELLS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MILLEN, WHITE, ZELANO & BRANIGAN, P.C.
; STREET: 2200 CLARENDON BLVD.
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22201

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/838,410
FILING DATE: 19920311
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP91/00942
FILING DATE: 15-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 185852-1990
FILING DATE: 13-JUL-1990
ATTORNEY/AGENT INFORMATION:
NAME: WHITE, JOHN L.
REGISTRATION NUMBER: 17,746
REFERENCE/DOCKET NUMBER: WAK 110
TELEPHONE: 703-243-6333
TELEFAX: 703-243-6410
TELEX: 64191
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2289 base pairs
TYPE: NUCLEIC ACID

QY 241 ccattcacttgcaaggcttttgttttgataaagcaagaaaaaactgcctctggtcccc 300
|||||
Db 241 CCATTCACTTGCAAGGCTTTGTGTTTGTATAAAGCAAGAAAAACAATGCCCTCTGGTCCCC 300
QY 301 ttcaatagcatgtcaagtggagtgaataaagaatttggccatgaatttgacctctatgaa 360
|||||
Db 301 TTCAATAGCATGTCAAGTGGAGTGAAAAAGAATTTGGCCATGAATTGACCTCTATGAA 360
QY 361 aacaaagactacattagaaactgcatcattggttaaaggacgcagctacaaagggaaacagta 420
|||||
Db 361 AACAAAGACTACATTAGAAACTGCATCATTTGGTAAAGGACGCAGCTACAAGGGAAACAGTA 420
QY 421 tctatcactaagagtggcatcaaatgtcagccctggagttccatgataccacacgaacac 480
|||||
Db 421 TCTATCACTAAGAGTGGCATCAAAATGTGAGCCCTGGAGTTCATGATACCACACGGAACAC 480
QY 481 agctatcggggtaaagacctacaggaataactactgtcgaatacctcgaggggaagagg 540
|||||
Db 481 AGCTATCGGGGTAAAGACCTACAGGAAACTACTGTGCAAAATCCTCGAGGGGAAGAGGG 540
QY 541 ggaccctggtgtttcacaaagcaatccagaggttacgtacgaagtctgtgacattcctcag 600
|||||
Db 541 GGACCCTGGTGTTCACAGCAATCCAGAGGTACGCTACGAAGTCTGTGACATTCCTCAG 600
QY 601 tgttcagaagtgaatgcatgacctgcaatggggagagttatcgagggtctcatggatcat 660
|||||
Db 601 TGTTCAGAAAGTTGAATGCATGACCTGCAATGGGGAGAGTTATCGAGGTCTCATGGATCAT 660
QY 661 acagaatcaggcaagatttgtcagcgtgggtatcatcagacacacaccgcgcacaaattc 720
|||||
Db 661 ACAGAATCAGGCAGGATTGTGACGCTGGGATCATCAGACACCACACCCGGCACAAATTC 720
QY 721 ttgcctgaaagatatcccgacaagggtttgtgatgataaattattgccgcaatccccgatggc 780
|||||
Db 721 TTGCCTGAAAGATATCCGACAAAGGCTTTTGATGATAATTATTGCCGCAATCCCGATGGC 780
QY 781 cagccgagggccatggtgtctatactcttgacccctcacaccgcgtgggagtactgtgcaatt 840
|||||
Db 781 CAGCCGAGGCCATGGTGTATACTCTTGACCCCTCACACCCGCTGGGAGTACTGTGCAATT 840
QY 841 aaaacatgcgctgacaaaagct 861
|||||
Db 841 AAAACATGCGCTGACAAATACT 861

RESULT 3

US-08-290-937B-4
; Sequence 4, Application US/08290937B

; Patent No. 5648233

; GENERAL INFORMATION:

; APPLICANT: YAMAGUCHI, KYOJI

; APPLICANT: SHIMA, NOBUYUKI

; APPLICANT: MURAKAMI, AKIHIKO

; APPLICANT: GOTO, MASAOKI

; APPLICANT: TSUDA, EISUKE

; APPLICANT: MASUNAGA, HIROAKI

; APPLICANT: TAKAHIRA, REIKO

; APPLICANT: OOGAKI, FUMIKO

; APPLICANT: UEDA, MASATSUGU

; APPLICANT: HIGASHIO, KANJI

; TITLE OF INVENTION: MODIFIED TCF

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Testa, Hurwitz & Thibault

; STREET: 125 High St.

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110

; COMPUTER READABLE FORM:

; - MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/290,937B

; FILING DATE: 19-AUG-1994

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: CAMPBELL, PAULA A.

; REGISTRATION NUMBER: 32,503

; REFERENCE/DOCKET NUMBER: FJN-022

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 248-7000

; TELEFAX: (617) 248-7100

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2288 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

US-08-290-937B-4

Query Match 49.0%; Score 845.8; DB 1; Length 2288;

Best Local Similarity 99.7%; Pred. No. 1.4e-216;

Matches 858; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1 atgtgggtgaccaaactcctgccagccctgctgctgctgctgctgctgctgctgctgctc 60

|||||

Db 78 ATGTGGGTGACCAAACTCCTGCCAGCCCTGCTGCTGCAGCATGTCTCTCTCTCTCCTC 137

Qy 61 ctgctcccatcgccatccctatgagaggggacaaaggaagaaatacaattcat 120

|||||

Db 138 CTGCTCCCATCGCCATCCCTATGCAGAGGGACAAAAGGAAAGAAATACAATTCTAT 197

Qy 121 gaattcaaaaaatcagcaaaagactaccctaatacaaaatagatccagcactgaagataaaa 180

|||||

Db 198 GAATTCAAAAAATCAGCAAAAGACTACCCCTAATCAAAA-AGATCCAGCAGTGAAGATAAAA 256

Qy 181 accaaaaaagtgaatactgagacccaatgtgctaataatagatgtactaggaataaaggactt 240

|||||

Db 257 ACCAAAAAAGTGAATACTGCAGACCAATGTGCTAATAGATGTACTAGGAATAAAGGACTT 316

Qy 241 ccattcacttgcaaggcttttgttttgataaaagcaagaaaaaactgcctctggttcccc 300

|||||

Db 317 CCATTCACTGCAAGGCTTTTGTGTTTGATAAAGCAAGAAACAATGCCTCTGTGTTCCCC 376

Qy 301 ttcaatagcatgtcaagtggagtgaataaagaaatttggccatgaatttgaccttatgaa 360

|||||

Db 377 TTCAATAGCATGTCAAGTGGAGTGAGTGAAAAAAGAAATTTGGCCATGAATTTGACCTCTATGAA 436

Qy 361 aacaaagactacattagaaaactgcatcattgttaaaggacgcagctacaaagggaaacagta 420

|||||

Db 437 AACAAAGACTACATTAGAAACTGCATCATTTGTTAAAGGACGCAGCTACAAGGGAAACAGTA 496

Qy 421 tctatcactaagagtggcatcaaatgtcagccctggagttccatgataccacacgaacac 480

|||||

Db 497 TCTATCACTAAGAGTGGCATCAAATGTGAGCCCTTGGAGTTCATGATACCACACGAACAC 556

Qy 481 agctatcggggtaaagacacctacaggaataactactgtcgaatacctcgaggggaagagg 540

|||||

Db 557 AGCTATCGGGGTAAAGACCTACAGGAAACTACTGTCGAAATCCTCGAGGGGGAAGAGGG 616

Qy 541 ggacctggtgtttcacaagcaatccagaggttacgctacgaagtctgtgacattcctcag 600

|||||

Db 617 GGACCCCTGGTGTTCACAAGCAATCCAGAGGTACGCTACGAAGTCTGTGACATTCCTCAG 676

Qy 601 tgttcagaagttgaatgcatgacctgcaatggggagaggttatcgagggtctcatgatcat 660

|||||

Db 677 TGTTCAGAAAGTTGAATGCATGACCTGCAATGGGAGAGGTTATCGAGGTCTCATGGATCAT 736

Qy 661 acagaatcaggcaagatttgtcagcgtgggatcatcagacaccacacccggcacaattc 720

|||||

Db 737 ACAGAATCAGGCAAGATTTGTGACGGCTGGGATCATCAGACACCACCCGGCCACAATTC 796


```

; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0912PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5416
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2232 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; PCT-US95-13830-1

Query Match          9.2%; Score 159; DB 5; Length 2232;
Best Local Similarity 59.1%; Pred. No. 3.3e-33;
Matches 295; Conservative 0; Mismatches 195; Indels 9; Gaps 1;

QY 346 ttgacctctatgaaacaaagactacattagaaactgcatcattggttaaaggacgcagc 405
Db 302 TGTGACCTCTTCCAGAAGAAAGACTACGTACGGACCTGCATCATGAACAATGGGTTGGG 361
QY 406 tacaagggaacagtatctatcactaagagtggcatcaaatgtcagccctggagttccatg 465
Db 362 TACCGGGGCACCATGGCCACAGACCGTGGTGGCTGCCCTGCCAGGCTTGGAGCCACAAG 421
QY 466 ataccacacgaacacagcta-----tcgggggttaaagacctacacaggaactactgt 516
Db 422 TTCCCGAATGATCACAAGTACACGCCCACTCTCCGGAATGGCCTGGAAGAGAACTTCTGC 481
QY 517 cgaatccctcaggggaagaagggggacccctggtgtttcacaaagcaatccagaggtacgc 576
Db 482 CGTAACCCCTGATGGCGACCCCGGAGGTCCCTTGGTGCTACACAACAGACCCTGCTGTGCGC 541
QY 577 tacgaagtctgtgacattcctcagtggttcagaagttgaatgcatgacctgcaatggggag 636
Db 542 TTCCAGAGCTGCGGCATCAAAATCCTGCCGGGAGGCCGCGTGTCTGGTGCAATGGCGAG 601
QY 637 agttatcgagggtctcatggtatcacagaatcagggaagatttgtcagcgtgggatcat 696
Db 602 GAATACCGCGCGCGGTAGACCGCAGCGAGTACGGGCGGAGTGCCAGCGCTGGGATCTT 661
QY 697 cagacaccacacccggcacaaattcttgccctgaaagatatcccgacaaaggcctttgatg 756
Db 662 CAGCACCCGACACACACCCCTTCGAGCCCGGCAAGTTCCTCGACCAAGGCTTGGACGAC 721
QY 757 aattattgcccgaatcccgatggccagccgagggccatggtgtatatactcttgacctcac 816
Db 722 AACTATTGCCGAATCCTGACGGCTCCGAGCGGCCATGGTGCTACACTACGGATCCGCAG 781
QY 817 acccgctgggagtactgtg 835
Db 782 ATCGAGCGAGAGTTCTGTG 800

RESULT 12
US-07-882-925A-7
; Sequence 7, Application US/07882925A
; Patent No. 5315000
; GENERAL INFORMATION:
; APPLICANT: Degen, Sandra J. F.
; TITLE OF INVENTION: Gene for a growth factor and its cDNA and
; TITLE OF INVENTION: protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gregory Lunn
; STREET: Wood, Herron & Evans, 2700 Carew Tower
; CITY: Cincinnati

```

```

; STATE: Ohio
; COUNTRY: USA
; ZIP: 45202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 6.0.3
; SOFTWARE: Microsoft Word 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/882,925A
; FILING DATE: 19920514
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Gregory
; REGISTRATION NUMBER: 29,945
; REFERENCE/DOCKET NUMBER: CMC 57
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 241-2324
; TELEFAX: (513) 421-7269
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2262 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; DESCRIPTION: Identical to sequence ID NO: 1: with 5' and 3'
; DESCRIPTION: adaptors added to make a full-length cDNA
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; DEVELOPMENTAL STAGE: fetal
; TISSUE TYPE: liver
; IMMEDIATE SOURCE:
; LIBRARY: cDNA
; CLONE: #33 including 5' and 3' adaptors
; PUBLICATION INFORMATION:
; RELEVANT RESIDUES IN SEQ ID NO: 7: FROM 1 TO 2262
; US-07-882-925A-7

Query Match          9.2%; Score 159; DB 1; Length 2262;
Best Local Similarity 59.1%; Pred. No. 3.3e-33;
Matches 295; Conservative 0; Mismatches 195; Indels 9; Gaps 1;

```

```

QY 346 ttgacctctatgaaacaaagactacattagaaactgcatcattggttaaaggacgcagc 405
Db 307 TGTGACCTCTTCCAGAAGAAAGACTACGTACGGACCTGCATCATGAACAATGGGTTGGG 366
QY 406 tacaagggaacagtatctatcactaagagtggcatcaaatgtcagccctggagttccatg 465
Db 367 TACCGGGGCACCATGGCCACAGACCGTGGTGGCTGCCCTGCCAGGCTTGGAGCCACAAG 426
QY 466 ataccacacgaacacagcta-----tcgggggttaaagacctacaggaactactgt 516
Db 427 TTCCCGAATGATCACAAGTACACGCCCACTCTCCGGAATGGCCTGGAAGAGAACTTCTGC 486
QY 517 cgaatccctcaggggaagaagggggacccctggtgtttcacaaagcaatccagaggtacgc 576
Db 487 CGTAACCCCTGATGGCGACCCCGGAGSTCCTTGGTGCTACACAACAGACCCTGCTGTGCGC 546
QY 577 tacgaagtctgtgacattcctcagtggttcagaagttgaatgcatgacctgcaatggggag 636
Db 547 TTCCAGAGCTGCGGCATCAAAATCCTGCCGGGAGGCCGCGTGTCTGTGTCGAATGGCGAG 606
QY 637 agttatcgagggtctcatggtatcacagaatcagggaagatttgtcagcgtgggatcat 696
Db 607 GAATACCGCGCGCGGTAGACCCGACCGAGTACAGGCGGCGGAGTGCACCGCTGGGATCTT 666
QY 697 cagacaccacacccggcacaaattcttgccctgaaagatatcccgacaaaggcctttgatg 756
Db 667 CAGCACCCGACACAGCACCCCTTCGAGCGGGGCAAGTTCCTCGACCAAGGCTTGGACGAC 726

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 16:21:42 ; Search time 2582.57 Seconds
(without alignments)
6313.932 Million cell updates/sec

Title: US-09-600-991-1
Perfect score: 1725
Sequence: 1 atgtgggtgaccaaactcct.....accaccaccaccaccactag 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	820.8	47.6	975	106	AL546558 AL546558
2	792.8	46.0	919	106	AL549164 AL549164
3	656.2	38.0	872	106	AL550803 AL550803
4	612	35.5	828	106	AL546883 AL546883
5	436	25.3	506	1	AA037738 zf03b09.r
6	432	25.0	844	106	AL546856 AL546856
7	389.2	22.6	899	106	AL573472 AL573472
8	270.2	15.7	412	110	AV759905 AV759905
9	166	9.6	484	1	AA037786 zf03b09.s
10	163.6	9.5	415	32	AV707491 AV707491
11	161	9.3	321	165	BE244598 TCBAPE08
12	136	7.9	661	112	AW175268 f134b04.y
13	135.8	7.9	753	227	AQ381428 RPCI11-13
14	132.8	7.7	766	148	BF383652 602044540
15	131.6	7.6	778	18	AI317783 u120f12.y
16	131.4	7.6	565	144	BF079866 230570 MA
17	128.2	7.4	518	139	BE755132 208867 MA
18	128	7.4	768	18	AI317745 u120b05.y
19	125.8	7.3	837	16	AI118965 ue93b02.y
20	124.4	7.2	844	21	AI530175 u190a05.y
21	122.8	7.1	470	232	AQ707814 HS_5561_B
22	122	7.1	470	166	BE326846 hr65g04.x
23	121.2	7.0	780	18	AI317794 u120h04.y
24	118.4	6.9	914	106	AL531543 AL531543
25	117.4	6.8	909	106	AL557971 AL557971
26	115.4	6.7	980	17	AI226536 u109f09.y
27	112.6	6.5	788	16	AI121047 u109f09.y
28	112.2	6.5	819	16	AI114957 u141a04.y
29	111.8	6.5	817	24	AI746861 u108d04.y
30	111	6.4	265	137	BE576686 dc35d11.y
31	110.8	6.4	773	16	AI119053 ue93g04.y
32	109.8	6.4	897	111	AW106317 um26b04.y
33	108	6.3	918	23	AI663558 u198c09.y
34	107.4	6.2	752	146	BF235733 602026881
35	107.4	6.2	780	146	BF234001 602024489
36	106.8	6.2	842	16	AI118969 ue93b06.y
37	105	6.1	910	21	AI530076 u188h06.y
38	102.2	5.9	889	106	AL558014 AL558014
39	101.4	5.9	794	24	AI746972 u109g09.y
40	99	5.7	741	16	AI119056 ue93g09.y
41	96.4	5.6	843	148	BF385803 602045966
42	94	5.4	243	248	AZ694675 AST-2HBG5
43	93.8	5.4	403	4	AA260053 va38b11.r
44	93	5.4	574	165	BE234370 141225 MA
45	93	5.4	594	104	AI958441 fd04g11.y

ALIGNMENTS

RESULT 1	AL546558	AL546558	975 bp	mRNA	EST	16-FEB-2001
LOCUS	AL546558	AL546558	LTI_NFL006_PL2	Homo sapiens	cdna clone	CS0DI030YH12 5
DEFINITION	prime, mRNA sequence.					
ACCESSION	AL546558					
VERSION	AL546558.1	GI:12879791				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 975)					
AUTHORS	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.					
TITLE	Full-length cdna libraries and normalization					
JOURNAL	Unpublished (2001)					
COMMENT	Contact: Genoscope					

Genoscope - Centre National de Sequencage	
BP 191 91006 EVRY cedex - France	
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.	
FEATURES	Location/Qualifiers
source	1. .975
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="CS0DI030YH12"
	/clone_lib="LTI_NFL006_PL2"
	/tissue_type="placenta"
	/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT	311 a 228 c 215 g 220 t 1 others
ORIGIN	
Query Match 47.6%; Score 820.8; DB 106; Length 975;	
Best Local Similarity 97.9%; Pred. No. 5.6e-211;	
Matches 858; Conservative 0; Mismatches 2; Indels 16; Gaps 2;	
QY	1 atgtgggtgacaaactcctgccagccctgctgctgcagcatgtcctcctcctcctc 60
Db	
QY	61 ctgctcccatcgccatccccctatgcagaggggacaaaagaaaagaaatacaattcat 120
Db	
QY	66 CTGCTCCCATCGCCATCCCTATGCGAGAGGACAAAAGGAAAGAAATACAATTCA 125
Db	
QY	121 gaattcaaaaaatcagcaaaagactaccctaatcaaaatagatccagcactgaagataaaa 180
Db	
QY	126 GAATTCAAAAAATCAGCAAGACTACCCCTAATCAAAATAGATCCAGCAGTAAAGATAAAA 185
Db	
QY	181 accaaaaagtgaatactactgcagaccatgtgctaataatagatgtactaggaataaaaggactt 240
Db	
QY	186 ACCAAAAAGTGAATACTGCAGACCAATGTCTAATAATAGATGTACTAGGAATAAAGGACTT 245
Db	
QY	241 ccattcaattgcaaggcttttgtttttataaaagcaaaaaaatgacctgtggttcccc 300
Db	
QY	246 CCATTCACTTCAAGGCTTTTGTGTTTATAAAGCAAGAAACAATGCCCTCTGTTTCCCC 305
Db	
QY	301 ttcaatagcatgtcaagtggagtgaataaaagaaatttggccatgaatttgaccttatgaa 360
Db	
QY	306 TTCAATAGCATGTCAAGTGGAGTGAAAAAAGAATTTGGCCATGAATTTGACCTCTATGAA 365
Db	
QY	361 aacaaagactacattagaaaaactgcatcattgggttaaaggacgcagctacaagggaacagta 420
Db	
QY	366 AACAAAGACTACATTAGAAACTGCATTCATTGGTAAAGGACGCAGCTACAAGGGAACAGTA 425
Db	
QY	421 tctatcaataagagtggcatcaaaatgtcagccctggagttccatgatataccacagaacac 480
Db	
QY	426 TCTATCACTAAGAGTGGCATCAAAATGTACGCCCTGGAGTTCCATGATACACACGAACAC 485
Db	
QY	481 -----agctatcgggggttaaagacctacaggaataactactgtcgaaatcct 525
Db	
QY	486 AGCTTTTTCCTTCGAGCTATCGGGGTAAAGACCTACAGGAAACTACTGTGGAATCCT 545
Db	
QY	526 cgaggggaagaaggggaccctggtgtttcacagaatccagaggttacgtacgaagtc 585
Db	
QY	546 CGAGGGGAAGAAGGGGACCCTGTTGTTTCAACAAGCAATCCAGAGGTACGCTACGAAGTC 605
Db	
QY	586 tgtgacattcctcagtggttcagaagttgaatgcatgacctgcaatggggagagttatcga 645
Db	
QY	606 TGTGACATTCTCAGTGTTCAGAAAGTTGAATGCATGACCTGCAATGGGGAGAGTTATCGA 665
Db	
QY	646 ggtctcatggatcatacacagaatcagggaagatttgtcagcgctggggtatcatcagacacca 705

/clone="CS0DI056YB22"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 274 a 212 c 110 g 266 t 10 others
ORIGIN

Query Match 38.0%; Score 656.2; DB 106; Length 872;
Best Local Similarity 89.3%; Pred. No. 1.7e-166;
Matches 735; Conservative 9; Mismatches 62; Indels 17; Gaps 3;
QY 1 atgtgggtgacaaaactcctgccagccctgctgctgcagcatgtcctcctgcctcctc 60
Db 51 ATTTTATTTTACAAAACTCCTTCCATCCCTTCTTCTTCATCATTTCTCCTCATCTCCTC 110
QY 61 ctgctccccatgccatccccctatgcagagggacaaaaggaaaagaatacaattcat 120
Db 111 CTCTCCCCATGCCATCCCCCTATTTCATATATGACAAATTAAATAATAATACAAATTCAT 170
QY 121 gaattcaaaaaatcagcaaaagactacccttaataataatagatccagcactgaagataaaa 180
Db 171 TAATTCAAAAATCATCAAAATACTACCCCTAATCAAAATATATCCATCATCAATAATAAAA 230
QY 181 accaaaaaagtgaatactgcagaccatgtgcttaataatagatgtactaggataaaaggactt 240
Db 231 ACCAAAAAATTAATACTGCATACCAATTTCTTAATATATTTACTATTATAATAATTAATT 290
QY 241 ccattcacttgcaaggctttgttttgataaaagcaagaaaaacaatgcctctgttcccc 300
Db 291 CCATTCACTTTCAAKKCTTTTTTTTTTATAAATCAAGAAAACAATTCCTCTGTTCCTCC 350
QY 301 ttcaatagcatgtcaagtggagtgaaaaagaatttggccatgaaatttgacctctatgaa 360
Db 351 TTCAATAGCATTTCAAGTTGAGTTAAAAAACAATTTTGCCATTAAATTTACCTCTATTAA 410
QY 361 acaaaagactacattagaaaactgcatcatgtggtaaaggcgcagctacaaagggaacagta 420
Db 411 AACAAAGACTACATTAKAACTTCATCATCTKGTAAAGGACGCTACAAAGGGAACAGTA 470
QY 421 tctatcactaagagtggcatcaaatgtcagccctggagttccatgataccacacgaacac 480
Db 471 TCTATCACTAAGAGTGGCATCAAAATGTCAGCCCTGGAGTTCCATGATACCACTAACAC 530
QY 481 -----agctatcggggttaaagacctacaggaataactactgtcgaaatcct 525
Db 531 ATCTTTTTCCTTCTATCTATCTKGGTAAAGACCTACAGGAAAACCTACTGTGGAATCCT 590
QY 526 cgaggggaagaggggggaccctgggttttcacaagcaatccagaggtacgctacgaagtc 585
Db 591 CTAYGTGAATAMGKGTGA-CCTGGTGTTCACAAGCAATCCAGAGGTACCTACGAAGTC 649
QY 586 tgt-gacattcctcagttgttcagaagttgaatgcatgacctgcaatgggggagagttatcg 644
Db 650 TGTGGACATTCCTCAGTGTTCAGAAGTTGAATGCATGACCTGCAATGGGGAGAGTTATCG 709
QY 645 aggtctcatggtatcatcacagaatcaggcaagatttgtcagcgctgggatcatcacagacc 704
Db 710 AGGTCTCATGGATCATACAGAATCAGGCAAGATTGTGACGCGTGGGATCATCAGACACC 769
QY 705 acaccggcacaataattcttgcctgaaagatatcccgacaaggccttggatgataattatg 764
Db 770 ACACCGGCACAAAATCTTTGCCTGAAAGATATCCCGACAAGGGCTTTGATGATAATTATTG 829

QY 765 ccgcaatcccgatggccagccgagggccatgggtgctatactctt 807
Db 830 CCGCAATCCCGATGGCCAGCGAGGCCCATGGTGCTATACTCTT 872
RESULT 4
AL546883 AL546883 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI026YD16 5
LOCUS prime, mRNA sequence.
DEFINITION AL546883 828 bp mRNA EST 16-FEB-2001
ACCESSION AL546883
VERSION AL546883.1 GI:12880433
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 828)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..828
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI026YD16"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 290 a 183 c 157 g 196 t 2 others
ORIGIN
Query Match 35.5%; Score 612; DB 106; Length 828;
Best Local Similarity 99.4%; Pred. No. 1.4e-154;
Matches 612; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 1 atgtgggtgacaaaactcctgccagccctgctgctgcagcatgtcctcctgcctcctc 60
Db 15 ATGTGGGTGACCAAACCTCTGCCAGCCCTGCTGCTGCAGCATGTCTCTCTCATCTCCTC 74
QY 61 ctgctccccatgcctatccccctatgcagagggacaaaaggaaaagaatacaattcat 120
Db 75 CTGCTCCCCATCGCCATCCCTATGCAGAGGGACAAAGGAAAGAAATAACAATTTCAT 134
QY 121 gaattcaaaaaatcagcaaaagactaccctaataataatagataatagtagaataaagactt 180
Db 135 GAATTCAAAAAATCAGCAAAAGACTACCCTAATCAAAATAGATCCAGCACTGAAGATAAAA 194
QY 181 accaaaaagtgaatactgcagaccatgtgctaataatagatgtactaggataaaggactt 240
Db 195 ACCAAAAAAGTGAATACTGCAGACCAATGTGCTAATAGATGTACTAGGAATAAAGGACTT 254
QY 241 ccattcacttgcaaggcttttgttttgataaaagcaagaaaaacaatgcctctggtcccc 300
Db 255 CCATTCACTTGCAGGGCTTTGTTTTTGATAAAGCAAGAAAACAATGCCTCTGTTCCCC 314
QY 301 ttcaatagcatgtcaagtggagtgaaaaaagaatttggccatgaatttgacctctatgaa 360
Db 315 TTCAATAGCATGTCAAGTGGAGTGAAAAAAGAATTTGGCCATGAATTTGACCTCTATGAA 374

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.
OM nucleic - nucleic search, using sw model
Run on: September 18, 2001, 17:44:38 ; Search time 4515.12 Seconds
(without alignments)
5796.398 Million cell updates/sec

Title: US-09-600-991-3
Perfect score: 1692
Sequence: 1 atgggggtggctccactcct.....accaccaccaccactag 1692

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_bal:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_bal:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vil:*
59: gb_vi2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rol:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES												
Result No.	Score	Query Match	Length	DB ID							Description	
1	1692	100.0	1692	9	AX019529	-265	97	HUMST1A	AX019529	Sequence	AX019529	Sequence
2	996.6	58.9	1725	9	AX019527	-265	10	I84668	AX019527	Sequence	AX019527	Sequence
3	829.2	49.0	2219	97			10	E12495		L11924 Homo sapien		
4	829.2	49.0	2232	10			10	E12495		I84668 Sequence 1		
5	827.6	48.9	2216	10			10	E12714		E12495 Human cDNA		
6	827.6	48.9	2216	10			97	HUMHEPGFA		E12714 Human hepat		
7	827.6	48.9	2216	97			10	I36332		M74178 Human hepat		
8	811.6	48.0	2219	10						I36332 Sequence 1		

I84668
LOCUS I84668 2232 bp DNA PAT 04-APR-1998
DEFINITION Sequence 1 from patent US 5696086.
ACCESSION I84668
VERSION I84668.1 GI:3022188
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2232)
AUTHORS Abraham,H.Karsenty and Godowski,P.J.
TITLE Methods and kits using macrophage stimulating protein
JOURNAL Patent: US 5696086-A 1 09-DEC-1997;
FEATURES
source Location/Qualifiers
1. .2232 /organism="unknown"
BASE COUNT 448 a 672 c 686 g 426 t
ORIGIN
Query Match 49.0%; Score 829.2; DB 10; Length 2232;
Best Local Similarity 99.6%; Pred. No. 1.4e-194;
Matches 831; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 atgggggtgctcccaactcctgctgcttctgactcaatgcttaggggtccctgggcagcgc 60
Db 11 ATGGGGTGGCTCCCACTCCTCTGCTGCTTCTGACTCAATGCTTAGGGGTCCCTGGGCAGCGC 70
QY 61 tcgccattgaatgacttccaagtctccgggggcacagagctacagcacctgctacatgcg 120
Db 71 TCGCCATTGAATGACTTCCAAGTGCTCCGGGGCACAGAGCTACAGCACCTGCTACATGCG 130
QY 121 gtggtgccccggcccttggcaggaggtgtggcagatgctgaagagtgtgctgctgcgtgt 180
Db 131 GTGGTGCCCGGCCCTTGGCAGGAGGATGTGGCAGATGCTGAAGAGTGCTGTGTCGCTGT 190
QY 181 gggcccttaatggactgcggggcccttcactacaacgtgagcagccatggttgccaaactg 240
Db 191 GGGCCCTTAATGGACTGCCGGGCCTTCCACTACAACGTGAGCAGCCATGGTTGCCAACTG 250
QY 241 ctgccatggactcaacactcgccccacacgaggctgcggcttctgggcgtgtgacctc 300
Db 251 CTGCCATGGACTCAACACTCGCCCCACACGAGGCTCGCGGCTTCTGGGCGCTGTGACCTC 310
QY 301 ttccagaagaagactacgtacggacctgcacatcatgaacaaatggggttgggtaccggggc 360
Db 311 TTCCAGAAGAAAGACTACGTACGGACCTGCATCATGAACAATGGGGTGGGTACCCGGGC 370
QY 361 accatggccacgacctgggtggcctgcccctgccaggcttgagccacaagtccccgaat 420
Db 371 ACCATGGCCACGACCGTGGTGGCCTGCCCTGCCAGGCTTGAGGCCACAAAGTTCCCGAAT 430
QY 421 gatcaagtatcacgccccactctccggaatggcctggaagagaacttctgccgtaaccct 480
Db 431 GATCACAAGTACACGCCCACTCTCCGGAATGGCCTGGAAGAGAACTTCTGCCGTAACCCCT 490
QY 481 gatggcgaccccgagggtccttggtgctacacacagaccctgctgtgcgcttccagagc 540
Db 491 GATGGGACCCCGGAGGTCCTTGGTGCTACACAACAGACCCCTGCTGTGCGCTTCCAGAGC 550
QY 541 tgcggcatcaaatcctgcgggagggccgcgtgtgtctgtgcaatggcgaggaataaccgc 600
Db 551 TCGGCATCAAAATCCTGCCGGAGGCGCGGTGTGTCTGGTGCAATGGCGAGGAATACCGC 610
QY 601 ggcgcgttagaccgcacggagtcagggcgcgagtgccagcgtgggatcttcagcaccccg 660
Db 611 GGC GCGGTAGACCGCACGGAGTCAGGGCGCGAGTGCCAGCGCTGGGATCTTCAGCACCCG 670
QY 661 caccagcacccttcgagcggggcaagtctcctcgaccaaggctctggacgacaactattgc 720
Db 671 CACCAACACCCCTTCGAGCCGGGCAAGTTCCTCGACCAAGGTCTGGACGACAATAATTGC 730
QY 721 cggaaatcctgacggctccgagcgggccatggtgtctacactacggatcccgagatcgagcga 780

Db 731 CGGAATCCTGACGGCTCCGAGCGCCATGCTGTCTACACTACGGATCCGCGAGATCGAGCGA 790
QY 781 gagttctgtgacctccccgcgtgcgggtccgagggcacagccccgcctcgagggc 834
Db 791 GAGTTCGTGACCTCCCCCGCTGGGGTCCGAGGCACAGCCCCGCCAAGAGGCC 844
RESULT 5
LOCUS E12495 2216 bp DNA PAT 24-JUN-1998
DEFINITION Human cDNA encoding macrophage stimulating protein.
ACCESSION E12495
VERSION E12495.1 GI:3251328
KEYWORDS JP 1997003099-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2216)
AUTHORS Yoshikawa,W., Shimonishi,M., Iwamoto,J., Takehara,T. and Hagiya,M.
TITLE MODIFIED MATERIAL OF MACROPHAGE STIMULATING PROTEIN AND ITS
JOURNAL Patent: JP 1997003099-A 1 07-JAN-1997;
COMMENT TOYOBO CO LTD
OS Homo sapiens (human)
PN JP 1997003099-A/1
PD 07-JAN-1997
PF 20-JUN-1995 JP 1995153309
PI YOSHIKAWA WATARU, SHIMONISHI MANABU, IWAMOTO JUNKO, PI
TAKEHARA TOYOHIRO,
PI HAGIYA MICHIO
PC C07K14/535,C07H21/04,C12N5/10,C12N15/09,C12P21/02,(C12N5/10,
PC C12R1:91),
PC (C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FH source 1. .2216 /organism='Homo sapiens'
FT /cell_line='Hep G2'
FT CDS 1. .2136 /product='macrophage stimulating protein' FT
FT sig_peptide 1. .54.
FEATURES
source Location/Qualifiers
1. .2216 /organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 445 a 666 c 682 g 423 t
ORIGIN
Query Match 48.9%; Score 827.6; DB 10; Length 2216;
Best Local Similarity 99.5%; Pred. No. 3.4e-194;
Matches 830; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 atgggggtgctcccaactcctgctgcttctgactcaatgcttaggggtccctgggcagcgc 60
Db 1 ATGGGGTGGCTCCCACTCCTGCTGCTTCTGACTCAATACTTAGGGGTCCCTGGGCAGCGC 60
QY 61 tcgccattgaatgacttccaagtctccgggggcacagagctacagcacctgctacatgcg 120
Db 61 TCGCCATTGAATGACTTCCAAGTGCTCCGGGGCACAGAGCTACAGCACCTGCTACATGCG 120
QY 121 gtggtgccccggcccttggcaggaggtgtggcagatgctgaagagtgtgctggtcgctgt 180
Db 121 GTGGTGCCCGGCCCTTGGCAGGAGGATGTGGCAGATGCTGAAGAGTGCTGTGTCGCTGT 180
QY 181 gggcccttaatggactgccgggccttcactacaacgtgagcagccatggttgccaaactg 240
Db 181 GGGCCCTTAATGGACTGCCGGGCCTTCCACTACAACGTGAGCAGCCATGGTTGCCAACTG 240

I36333	I36333	2219 bp	DNA	PAT	13-MAY-1997
LOCUS	Sequence 2 from patent US 5606029.				
DEFINITION	I36333				
ACCESSION	I36333.1				
VERSION	GI:2086846				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 2219)				
AUTHORS	Degen,S.J.				
TITLE	Gene for a growth factor and its cDNA and protein				
JOURNAL	Patent: US 5606029-A 2 25-FEB-1997;				
FEATURES	Location/Qualifiers				
source	1. .2219				
BASE COUNT	461 a	660 c	674 g	420 t	4 others
ORIGIN					
Query Match	47.9%; Score 810; DB 10; Length 2219;				
Best Local Similarity	99.4%; Pred. No. 7.7e-190;				
Matches 813; Conservative	0;	Mismatches	5;	Indels	0; Gaps 0;
QY	17	tcctgctgtctgactcaatgcttaggggtccctggcagcgctgccattgaatgact	76		
Db	1	TCCTGCTGCTTCTGACTCAATACTTAGGGGTCCCCTGGCAGCGCTCGCCATTGAATGACT	60		
QY	77	tcaaagtgtccggggcacagagctacagcacctgtacatgcggtgtgcccgggcctt	136		
Db	61	TCCAAAGTGCTCCGGGGCACAGAGCTACAGCACCTGTGTACATGCGGTGGTGC	120		
QY	137	ggcaggaggtgtggcagatgctgaagagtgtgctggtgcgtgtggcccttaatggact	196		
Db	121	GGCAGGAGGATGTGGCAGATGCTGAAGAGTGTGCTGGTGCCTTAAATGGACT	180		
QY	197	gccgggcttccactacaacgtgagcagccatggttgccaaactgctgccatggactcaac	256		
Db	181	GCCGGGCTTCCACTACAACTGAGCAGCCATGGTTGCCAACTGCTGCCATGGACTCAAC	240		
QY	257	actgcggccacacagagctgcggcgttctggtggcgctgtgacctcttccagaagaagact	316		
Db	241	ACTCGCCCCACACGAGGCTGCGGCGTTCTGGGCGCTGTGACCTCTTCCAGAAGAAAGACT	300		
QY	317	acgtacggacctgcatacgaacaaatggggttggttacccggggcaccatggccacgaccg	376		
Db	301	ACGTACGGACCTGCATCATGAACAATGGGGTTGGGTACCGGGGCACCATGGCCACGACCG	360		
QY	377	tgggtggcctgccctgccaggttgagccacaagtctcccgaatgatcacaagtacacgc	436		
Db	361	TGGGTGGCCTGCCCTGCCAGGCTTGAGCCACAAAGTTCCCGAATGATCACAAAGTACACGC	420		
QY	437	ccactctcgggaatggcctggaagagaacttctgcgtaaccctgatggcgacccccggag	496		
Db	421	CCACTCTCCGGAATGGCCTGGAAGAGAACTTTCGCCGTAACCCCTGATGGCGACCCCGGAG	480		
QY	497	gtccttgggtgtacacaacagaccctgctgtgcgttccagagctgcggcatcaaatcct	556		
Db	481	GTCCTTGGTGTACACAACAGACCCCTGCTGTGCGCTTCCAGAGCTGCGGCATCAAAATCCT	540		
QY	557	gccgggagccgcgctgtgtctggtgcaatggcgaggaataaccggcgcgcggttagaccgca	616		
Db	541	GCCGGGAGGCCGCGTGTGTCTGGTGAATGGCGAGGAATAACCGGCGCGGTAGACCGCA	600		
QY	617	cggagtcaggcgcgagtgccagcgctgggatcttcagcacccgcaccagcacccttcg	676		
Db	601	CGGAGTCAGGGCGCGAGTTCACGCGCTGGGATCTTCAGCACCCCGCACCGACCCCTTCG	660		
QY	677	agccgggcaagtctcctcgaccaaggctctggacgacaaactattgccgggaatcctgacggct	736		
Db	661	AGCCGGCAAGTTCCTCGACCAAGGTCTGGACGCAAACTATTGCCGGGAATCCTGACGGCT	720		
QY	737	ccgagcggccatggtgtctacactacggaatccgcagatcgagcgagagttctgtgacctcc	796		

Db	721																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
----	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 18:33:25 ; Search time 214.84 Seconds
(without alignments)
4945.120 Million cell updates/sec

Title: US-09-600-991-3
Perfect score: 1692
Sequence: 1 atgggggtgctccactcct.....accaccaccaccactag 1692

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601:*
1: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1692	100.0	1692	20	AAX87663
2	996.6	58.9	1725	20	AAX87662
3	827.6	48.9	2213	18	AAT47149
4	827.6	48.9	2216	18	AAT47145
5	827.6	48.9	2216	20	AAV72083
6	811.6	48.0	2219	15	AAQ79723
7	811.6	48.0	2219	18	AAT62436
8	811.6	48.0	2262	15	AAQ79729
9	811.6	48.0	2262	18	AAT62439
10	810	47.9	2219	15	AAQ79724
11	810	47.9	2219	18	AAT62437

12	794.2	46.9	2021	18	AAT62438	Human L5/3 partial
13	792.6	46.8	2021	15	AAQ79725	Human L5/3 tumour
14	773	45.7	2190	20	AAX87676	HGF-MSP hybrid pro
15	773	45.7	2190	20	AAX87677	HGF-MSP hybrid pro.
16	773	45.7	2199	12	AAQ15177	Human leukocyte-de
17	773	45.7	2199	19	AAV53627	Human leukocyte-de
18	773	45.7	2289	11	AAQ06088	Tumour cytotoxic f
19	771.4	45.6	2172	19	AAV61952	Nucleotide sequenc
20	771.4	45.6	2173	14	AAQ37308	Encodes haematopoi
21	771.4	45.6	2289	13	AAQ21066	TCF II-encoding se
22	766.8	45.3	2172	12	AAQ14182	Plasminogen-like g
23	748	44.2	2184	13	AAQ20049	Human hepatocyte g
24	748	44.2	2184	14	AAQ46040	Hepatocyte growth
25	748	44.2	2187	14	AAQ47832	Competative inhibi
26	748	44.2	2187	14	AAQ47833	Competative inhibi
27	748	44.2	2187	14	AAQ45702	Vascular endotheli
28	748	44.2	2214	12	AAQ15176	Human leukocyte-de
29	748	44.2	2214	19	AAV53626	Human leukocyte-de
30	746.4	44.1	2184	15	AAQ56157	Hepatocyte growth
31	746.4	44.1	2187	12	AAQ10489	Hepatic parenchyma
32	746.4	44.1	2393	12	AAQ14038	Human hepatocyte g
33	743.2	43.9	2184	13	AAQ26052	Human HGF. Homo s
34	739.2	43.7	873	13	AAQ23662	Truncated hepatocy
35	739.2	43.7	873	14	AAQ34613	cDNA encoding 34kD
36	739.2	43.7	873	18	AAT78409	34 kilodalton hepa
37	735.8	43.5	1661	12	AAQ12398	Human hepatocyte g
38	735.8	43.5	1661	13	AAQ22144	Human HGF gene par
39	735.8	43.5	2187	12	AAQ12399	Human hepatocyte g
40	735.8	43.5	2187	13	AAQ22146	Complete human HGF
41	735.8	43.5	2187	19	AAV20546	Human recombinant
42	735.8	43.5	2187	20	AAV15186	Nucleic acid encod
43	615.8	36.4	1400	13	AAQ20745	Encodes alpha-chai
44	614.2	36.3	2187	13	AAQ20255	Encodes alpha- and
45	564.2	33.3	2216	20	AAV72084	Mouse MSP cDNA. M

ALIGNMENTS

RESULT	1
AAX87663	
ID	AAX87663 standard; cDNA; 1692 BP.
XX	
AC	AAX87663;
XX	
DT	26-OCT-1999 (first entry)
XX	
DE	HGF-MSP hybrid protein Metron Factor-1 cDNA.
XX	
KW	Metron Factor-1; human; hepatocyte growth factor; HGF;
KW	macrophage stimulating proteinl MSP; apoptosis; chemotherapy;
KW	toxicity; therapy; ss.
XX	
OS	Homo sapiens.
OS	Synthetic.
XX	
FH	Key
FT	sig_peptide
FT	mat_peptide
FT	
XX	
PN	WO9938967-A2.
XX	
PD	05-AUG-1999.
XX	
PF	27-JAN-1999; 99WO-EP00478.
XX	
PR	30-JAN-1998; 98IT-MI00179.
XX	
PA	(DOMP-) DOMPE SPA.
XX	
PI	Caselli G, Collesi C, Comoglio P, Medico E, Michieli P;

QY 1561 aatcccgatgccagcgagcccatggtgctatactcttgaccctcacacccgctgggag 1620
|||||
Db 1561 aatcccgatgccagcgagcccatggtgctatactcttgaccctcacacccgctgggag 1620
QY 1621 tactgtgcaattaaacatgcgctgacaaaagctgacgacgacgacacacaccaccac 1680
|||||
Db 1621 tactgtgcaattaaacatgcgctgacaaaagctgacgacgacgacacacaccaccac 1680
QY 1681 caccaccactag 1692
|||||
Db 1681 caccaccactag 1692

RESULT 2
AAx87662

ID AAX87662 standard; cDNA; 1725 BP.

XX AAX87662;

DT 26-OCT-1999 (first entry)

DE Hepatocyte growth factor hybrid protein Magic F-1 cDNA.

XX Magic Factor-1; human; hepatocyte growth factor; HGF; apoptosis;
KW chemotherapy; toxicity; therapy; ss.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT sig_peptide 1..63

FT /*tag= a

FT mat_peptide 64..1725

FT /*tag= b

XX WO9938967-A2.

PN 05-AUG-1999.

XX 27-JAN-1999; 99WO-EP00478.

PR 30-JAN-1998; 98IT-MI00179.

XX (DOMP-) DOMPE SPA.

PI Caselli G, Collesi C, Comoglio P, Medico E, Michieli P;

DR WPI; 1999-494090/41.

DR P-PSDB; AAY06619.

XX Recombinant proteins from recombination of HGF and MSP structural
PT domains, useful for protection against apoptosis induced by
PT chemotherapeutics

XX Example 2; Page 56-57; 63pp; English.

XX This DNA sequence codes for Magic F-1 factor (see AAY06619), a
CC recombinant protein composed of the hepatocyte growth factor (HGF)
CC alpha chain signal peptide, hairpin loop and kringle domains 1 and
CC 2 joined via a peptide linker to the HGF hairpin loop and kringle
CC domains 1 and 2 and a polyhistidine tag. The portions of DNA
CC encoding the various regions of Magic F-1 were obtained by PCR
CC amplification and then recombined to obtain the hybrid sequence.
CC Expression vectors, prokaryotic or eukaryotic host cells and a
CC process for preparing recombinant proteins from HGF and MSP are
CC claimed. Magic F-1 and other recombinant proteins comprising HGF
CC and/or macrophage stimulating protein (MSP) structural domains are
CC used to prevent or treat chemotherapeutic-induced toxicity such as
CC myelotoxicity, hepatotoxicity, nephrotoxicity, mucotoxicity and
CC neurotoxicity (claimed). They protect cells from death (apoptosis)
CC induced by chemotherapy of, e.g. tumors. In particular, they can
CC be used for expansion of marrow precursors, to increase
CC proliferation of the haematopoietic precursors or to stimulate

CC their entry in the cycle. By modification of the proteolytic
CC site, hybrid factors can be obtained which are activated by
CC proteases of the endoplasmic reticulum (such as furines) during
CC their synthesis. When the proteolytic site is removed, permanently
CC immature forms of the factors can be obtained, having a potential
CC partial agonistic or antagonistic activity. Different functional
CC domains can be combined so as to modulate the biological effects.
XX
SQ Sequence 1725 BP; 563 A; 399 C; 390 G; 373 T; 0 other;

Query Match 58.9%; Score 996.6; DB 20; Length 1725;
Best Local Similarity 83.2%; Pred. No. 1.7e-237;
Matches 1166; Conservative 0; Mismatches 214; Indels 21; Gaps 2;

QY 292 tgtgacctcttccagaagaagactacgtacggacctgcatcatgaacaatggggttggg 351
|||||
Db 346 ttgacctctatgaaaacaaagactacattagaaactgcatcattgtaaaggacgcgc 405
QY 352 tacgggggcaccatggccacgacogtgggtggcctgccctggcaggctggagccacaag 411
|||||
Db 406 tacaagggaacagtatctatcactaaagtggtgcataaaatgtcacccctggagttccatg 465
QY 412 ttccggaatgatcacaaagtacacgccactctccggaatggcctgggaagagaacttctgc 471
|||||
Db 466 ataccacacgaacacagcta-----tcggggtaaaagacctacaggaaaaactactgt 516
QY 472 cgtaacacctgatggcgaccccgagggttccttgggtgtctacacacagaccctgctgtgcgc 531
|||||
Db 517 cgaatacctcgagggggaagaagggggaccctggtttcacaaagaatccagaggtacgc 576
QY 532 ttccagagctgcggcatcaaatctcctccgggagccgctgtgtctgtgcaatggcgag 591
|||||
Db 577 tacgaagtctgtgacattcctcagttgttcagaagtgaatgcattgacctgcaatggggag 636
QY 592 gaataccgcgcgcggttagaccgcagggagtcaggcgcgagtgccagcgcctgggatctt 651
|||||
Db 637 agttatcgaggtctcatggtacatacagaatcaggcaagattgtcagcgcctgggatcat 696
QY 652 cagcaccgcgacaccctctcgagccgggcaagtctcctcgaccaaggctctggacgac 711
|||||
Db 697 cagacaccacacccggcacaaattcttgctgaaagatatcccgacaaaggcctttgatgat 756
QY 712 aactattgcgggaatcctgacggtccgagcgccatggtgctacactacggatccgcag 771
|||||
Db 757 aattattgccgcaatcccgatggccagccgagggccatggtgctatactcttgaccctcac 816
QY 772 atcgagcgagagttctgtgacctccccctgcggttcgagggcacagccccgcctcgag 831
|||||
Db 817 accgctgggagtactgtgcaattaaacatgcgctgacaaagc-----ttcg 864
QY 832 ggcggtggcggttctggtggcggtggttcggttcggtgcggttctctagagggacaaaag 891
|||||
Db 865 ggcggtggcggttctggtggcggtggttcggttcggttcctctagagggacaaaag 924
QY 892 aaaagaagaaatacaattcatgaattcaaaaaaatcagcaaaagactaccctaatcaaaata 951
|||||
Db 925 aaaagaagaaatacaattcatgaattcaaaaaaatcagcaaaagactaccctaatcaaaata 984
QY 952 gatccagcactgaagataaaaccccaaaagtgaatactgcagaccaatgtgctaataga 1011
|||||
Db 985 gatccagcactgaagataaaaccccaaaagtgaatactgcagaccaatgtgctaataga 1044
QY 1012 tgtactaggaataaaaggacttccattcacttgcaggcctttgttttgataaaagcaaga 1071
|||||
Db 1045 tgtactaggaataaaaggacttccattcacttgcaggcctttgttttgataaaagcaaga 1104
QY 1072 aaacaatgcctctggttcccccttcaatagcatgtcaagtggagtgaaaaagaatttggc 1131
|||||
Db 1105 aaacaatgcctctggttcccccttcaatagcatgtcaagtggagtgaaaaagaatttggc 1164
QY 1132 catgaatttgacctctatgaaaacaaagactacattagaaaactgcatcattgggtaaagga 1191
|||||

Db 661 caccagcaccctcgagccgggcaagtctctcgaccaaggtctggacgacaactattgc 720
QY 721 cggaatcctgacggtccgagcggcccatggtgtctacactacggtacccgagatcgagcgca 780
Db 721 cggaatcctgacggtccgagcggcccatggtgtctacactacggtacccgagatcgagcgca 780
QY 781 gagttctgtgacctcccccgctgcggggtccgagggcacagccccgcctcgaggggc 834
Db 781 gagttctgtgacctcccccgctgcggggtccgagggcacagccccgcgaagaggcc 834

RESULT 4
AAT47145
ID AAT47145 standard; cDNA; 2216 BP.
XX
AC AAT47145;
DT 27-MAR-1997 (first entry)
XX
DE Macrophage stimulating protein C672X variant cDNA.
XX
KW Macrophage stimulating protein; MSP; protein engineering;
KW disulphide bond; ds.
XX
OS Synthetic.
FH Key Location/Qualifiers
FT CDS 1..2136 /*tag= a
FT sig_peptide 1..54 /*tag= b
FT mat_peptide 55..2133 /*tag= c
FT misc_difference 2014..2016 /*tag= d
FT /*note= "nnn at positions 2014-2016 represent an
FT alternative codon to the native cysteine
FT (tgc) codon, partic. GCN encoding alanine"

PN EP750040-A2.
XX
PD 27-DEC-1996.
XX
PF 18-JUN-1996; 96EP-0109750.
XX
PR 20-JUN-1995; 95JP-0153309.
XX
PA (TOYM) TOYO BOSEKI KK.
PA (TOYM) TOYOBO KK.
XX
PI Hagiya M, Iwamoto J, Shimonishi M, Takehara T, Yoshikawa W;
XX
DR WPI; 1997-044835/05.
DR P-PSDB; AAW07692.
XX
PT Macrophage stimulating protein mutant - with cysteine deletion or
PT substitution, can be produced recombinantly with minimal loss of
PT activity due to incorrect disulphide bond formation
XX
PS Claim 11; Page 12-14; 21pp; English.

XX
CC A DNA fragment (AAT47145) encodes a human macrophage stimulating
CC protein (MSP) mutant (AAW07692) in which the native cysteine residue
CC at position 672 is replaced by another amino acid, pref. alanine.
CC The native gene is subjected to site-directed mutagenesis to alter
CC codon 672. Alternatively, the Cys672 codon is deleted (see also
CC AAT47149). The mutant MSPs can be mass produced in transformed host
CC cells with minimal loss of activity due to incorrect disulphide
CC bond formation.
XX
SQ Sequence 2216 BP; 445 A; 665 C; 681 G; 422 T; 3 other;

Query Match 48.9%; Score 827.6; DB 18; Length 2216;
Best Local Similarity 99.5%; Pred. No. 1.5e-195;
Matches 830; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 atgggggtggtcccaactcctgctgtcttctgactcaatgcttaggggtccctggcagcgc 60
Db 1 atgggggtggtcccaactcctgctgtcttctgactcaatacttaggggtccctggcagcgc 60
QY 61 tcgccattgaatgacttccaagtgtcccggtccgggcacagagctacagacctgtacatgcg 120
Db 61 tcgccattgaatgacttccaagtgtcccggtccgggcacagagctacagacctgtacatgcg 120
QY 121 gtggtgccccggccttggcaggagatgtggcagatgctgaagagtgtgctgctgctgt 180
Db 121 gtggtgccccggccttggcaggagatgtggcagatgctgaagagtgtgctgctgctgt 180
QY 181 gggcccttaatggactgcggggtcccttccactacaacgtgagcagccatggttgcgaactg 240
Db 181 gggcccttaatggactgcggggtcccttccactacaacgtgagcagccatggttgcgaactg 240
QY 241 ctgccatggactcaaacactgcgcccaacagaggtgcgggttcttgggcgctgtgacctc 300
Db 241 ctgccatggactcaaacactgcgcccaacagaggtgcgggttcttgggcgctgtgacctc 300
QY 301 ttccagaagaagactactgtacggacctgcctgccctgccaggttggagccacaagtccccgaat 360
Db 301 ttccagaagaagactactgtacggacctgcctgccctgccaggttggagccacaagtccccgaat 360
QY 361 accatggccacgacctggtgggtggtccctgccctgccaggttggagccacaagtccccgaat 420
Db 361 accatggccacgacctggtgggtggtccctgccctgccaggttggagccacaagtccccgaat 420
QY 421 gatcacaagtacacgcccactctccgggaatggcctggaaagagaacttctgccgttaaccct 480
Db 421 gatcacaagtacacgcccactctccgggaatggcctggaaagagaacttctgccgttaaccct 480
QY 481 gatggcgacccccggaggtccttgggtctacacaaacagacctgctgtgcgcttccagagc 540
Db 481 gatggcgacccccggaggtccttgggtctacacaaacagacctgctgtgcgcttccagagc 540
QY 541 tgcggcatcaaatcctgcggggagcgcggtgtgtctgtgcaatggcgagggaataccgc 600
Db 541 tgcggcatcaaatcctgcggggagcgcggtgtgtctgtgcaatggcgagggaataccgc 600
QY 601 ggcgcggttagaccgcacgagtcagggtcgaggtccagcgtggtccagcgttccagcaccgc 660
Db 601 ggcgcggttagaccgcacgagtcagggtcgaggtccagcgtggtccagcgttccagcaccgc 660
QY 661 caccagcacccttcgagcggggaagtccctcgaccagggtctggacgacaactattgc 720
Db 661 caccagcacccttcgagcggggaagtccctcgaccagggtctggacgacaactattgc 720
QY 721 cggaatcctgacgggtccgagcgggccaatggtgtctacactacggtcgcgagatcgagcgca 780
Db 721 cggaatcctgacgggtccgagcgggccaatggtgtctacactacggtcgcgagatcgagcgca 780
QY 781 gagttctgtgacctcccccgctgcgggtccgagggcacagccccgcctcgaggggc 834
Db 781 gagttctgtgacctcccccgctgcgggtccgagggcacagccccgcgaagaggcc 834

RESULT 5
AAV72083
ID AAV72083 standard; cDNA; 2216 BP.
XX
AC AAV72083;
XX 12-APR-1999 (first entry)
XX Human MSP cDNA.
XX MSP; macrophage stimulating protein; apoptosis; human; treatment;
KW neuroendocrine cell; RON receptor; small cell lung carcinoma; tumour;


```
Db 152 gccaggaggatgtggcagatgctgaagagtgtgctgctggtgctgtgggcccttaatggact 211
Qy 197 gccgggacctccactacaacgtgagcagccatggttgccaactgctgccatggactcaac 256
Db 212 gccgggacctccactacaacgtgagcagccatggttgccaactgctgccatggactcaac 271
Qy 257 actegcccccacacgaggctggcggttcttgggcgctgtgacctcttccagaagaagact 316
Db 272 actcgccccacacgaggctgcggcttcttgggcgctgtgacctcttccagaagaagact 331
Qy 317 acgtacggacctgcacatgaacaatatggggttggttacccggggcaccatggccacgaccg 376
Db 332 acgtacggacctgcacatgaacaatatggggttggttacccggggcaccatggccacgaccg 391
Qy 377 tgggtggcctgccctgccaggcttgagccacaagtcccgaaatgatcacaaagtacacgc 436
Db 392 tgggtggcctgccctgccaggcttgagccacaagtcccgaaatgatcacaaagtacacgc 451
Qy 437 ccactctccggaatggcctggaagagaacttctgccgttaacctgatggcacccccggag 496
Db 452 ccactctccggaatggcctggaagagaacttctgccgttaacctgatggcacccccggag 511
Qy 497 gtccttggtgtacacaaacagaccctgctgtgctgcttccagagctgcggcatcaaatcct 556
Db 512 gtccttggtgtacacaaacagaccctgctgtgcttccagagctgcggcatcaaatcct 571
Qy 557 gccgggagccgcgtgtgtctggtgcaatggcgaggagaataccgcggcggtagaccgca 616
Db 572 gccgggagccgcgtgtgtctggtgcaatggcgaggagaataccgcggcggtagaccgca 631
Qy 617 cggagtcagggcgcgagtgccagcgcctgggatcttgacgacaaactattgccggaatcctgacggct 676
Db 632 cggagtcagggcgcgagtgccagcgcctgggatcttgacgacaaactattgccggaatcctgacggct 691
Qy 677 agccgggcaagtctcctcgaccaaggtctggacgacaaactattgccggaatcctgacggct 736
Db 692 agccgggcaagtctcctcgaccaaggtctggacgacaaactattgccggaatcctgacggct 751
Qy 737 ccgagcggccatggtgtctacctacggtacggtccgcagatccgagcagagattctgtgacctcc 796
Db 752 ccgagcggccatggtgtctacctacggtacggtccgcagatccgagcagagattctgtgacctcc 811
Qy 797 ccgctgcgggtccgagggcacagccccgcctcgagggc 834
Db 812 ccgctgcgggtccgagggcacagccccgcgaaggcc 849

RESULT 9
AAT62439
ID AAT62439 standard; cDNA to mRNA; 2262 BP.
XX
AC AAT62439;
XX
DT 18-JUL-1997 (first entry)
XX
DE Human L5/3 construct gene.
XX
KW Human; growth factor; foetal; liver; probe; bovine; prothrombin; locus;
KW polymorphism; transition; exon; intron; chromosome; kringle domain; ss;
KW cell growth; tumour suppressor; hepatocyte growth factor; regeneration.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_signal 5..13
FT FT /*tag= a
FT FT /note= "consensus Kozak translation initiation sequence"
FT CDS 10..2151
FT FT /*tag= b
FT FT /product= human L5/3 protein
FT sig_peptide 10..108
FT FT /*tag= c
FT mat_peptide 109..2148
```

```
FT /*tag= d
XX
PN US5606029-A.
XX
PD 25-FEB-1997.
XX
PF 14-MAY-1992; 92US-0882925.
XX
PR 14-MAY-1992; 92US-0882925.
PR 18-JAN-1994; 94US-0184012.
XX
PA (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX
PI Degen SJ;
XX
DR WPI; 1997-153621/14.
DR P-PSDB; AAW14269.
XX
PT Human growth factor protein L5/3 - useful for altering cell growth,
PT e.g. as tumour suppressor
XX
PS Disclosure; Column 53-58; 34pp; English.
XX
CC This is the nucleotide sequence of a construct which encodes the human
CC growth factor designated L5/3 (AAW14270). This sequence was generated
CC using the clone #33 sequence (AAT62436) with the addition of adapters at
CC the 5' and 3' ends. The 3' adapter does not alter the encoded protein
CC sequence as it is inserted in the untranslated region. However the 5'
CC adapter was used to reconstruct the 5' coding sequence. The adapter
CC contained a 5' EcoRI restriction enzyme site followed by a Kozak
CC translation initiation consensus sequence. At the 3' end of the
CC adapter, an overhang compatible with the EcoRI site of the clone #33 cDNA
CC was used for ligation. Ligation of the adapter destroyed the EcoRI site
CC from the clone #33 sequence leaving only the adapter's EcoRI site for
CC cloning. Destruction of the clone #33 EcoRI site also resulted in a
CC substitution of the encoded Leu at position 6 of the wild type protein
CC with a tripeptide Asn-Ser-Val in this construct. The protein can be used
CC to alter cell growth (as a growth factor or tumour suppressor) and has
CC similar properties to the hepatocyte growth factor that is actively
CC involved in liver regeneration.
XX
SQ Sequence 2262 BP; 468 A; 676 C; 689 G; 429 T; 0 other;

Query Match 48.0%; Score 811.6; DB 18; Length 2262;
Best Local Similarity 99.5%; Pred. NO. 1.4e-191;
Matches 814; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 17 tcctgctcttctgactcaatgcttaggggtccctgggcagcgcctgccaatgaatgact 76
Db 32 tcctgctcttctgactcaatacttaggggtccctgggcagcgcctgccaatgaatgact 91
Qy 77 tccaagtgtccgggggcacagagctacagcacctgtctacatgcggtggtgcccgccctt 136
Db 92 tccaagtgtccgggggcacagagctacagcacctgtctacatgcggtggtgcccgccctt 151
Qy 137 gccaggaggatgtggcagatgctgaagagtgtgctggtcgctgtggcccttaatggact 196
Db 152 gccaggaggatgtggcagatgctgaagagtgtgctggtcgctgtggcccttaatggact 211
Qy 197 gccgggacctccactacaacgtgagcagccatggttgccaaactgctgccatggactcaac 256
Db 212 gccgggacctccactacaacgtgagcagccatggttgccaaactgctgccatggactcaac 271
Qy 257 actcgccccacacgaggctgcggcggttcttgggcgctgtgacctcttccagaagaagact 316
Db 272 actcgccccacacgaggctgcggcggttcttgggcgctgtgacctcttccagaagaagact 331
Qy 317 acgtacggacctgcacatgaacaatatggggttggttacccggggcaccatggccacgaccg 376
Db 332 acgtacggacctgcacatgaacaatatggggttggttacccggggcaccatggccacgaccg 391
Qy 377 tgggtggcctgccctgccaggcttgagccacaagtcccgaaatgatcacaaagtacacgc 436
```

Db 661 agcgggcaagtctcctgaccaaggtctgtggacgacaaactattgccgggaatcctgacggct 720
QY 737 ccgagcgccatggtgtacactacggtacggtccgagatcgagcgagagttctgtgacctcc 796
Db 721 ccgagcgccatggtgtacactacggtacggtccgagatcgagcgagagttctgtgacctcc 780
QY 797 ccgctgcgggtccgagggcacagccccgcctcgagggc 834
Db 781 ccgctgcgggtccgagggcacagccccgcctcgaggcc 818

RESULT 11
AAT62437
ID AAT62437 standard; cDNA to mRNA; 2219 BP.
XX AC AAT62437;
XX DT 18-JUL-1997 (first entry)
XX DE Human L5/3 partial clone #33 polymorphism #2.
XX KW Human; growth factor; foetal; liver; probe; bovine; prothrombin; locus;
KW polymorphism; transition; exon; intron; chromosome; kringle domain; ss;
KW cell growth; tumour suppressor; hepatocyte growth factor; regeneration.
XX OS Homo sapiens.

XX FH Key Location/Qualifiers
FT CDS 3..2120
FT /tag= a
FT /product= part of human L5/3 protein
FT sig_peptide 3..77
FT /tag= b
FT mat_peptide 78..2117
FT /tag= c
FT variation replace(618..620, TGC)
FT /tag= d
XX

PN US5606029-A.
XX
XX 25-FEB-1997.
PD
XX
PF 14-MAY-1992; 92US-0882925.
XX
XX 14-MAY-1992; 92US-0882925.
PR
PR 18-JAN-1994; 94US-0184012.
XX
XX (CHIL-) CHILDREN'S HOSPITAL MEDICAL CENT.
XX PA
XX Degen SJ;
PI
XX WPI; 1997-153621/14.
DR P-PSDB; AAW14267.
XX

PT Human growth factor protein L5/3 - useful for altering cell growth,
PT e.g. as tumour suppressor
XX
XX Disclosure; Column 17-24; 34pp; English.
PS
XX This is the nucleotide sequence of the insert isolated from clone #33
CC and encodes part of a human growth factor designated L5/3 (AAW14270).
CC This sequence was isolated from a human foetal liver cDNA library using
CC a fragment of the corresponding genomic sequence as a probe. This
CC fragment was isolated from a human liver genomic DNA library using a 1200
CC bp probe derived from the bovine prothrombin gene. The longest human cDNA
CC isolated contained this sequence and a polymorphic sequence (AAT62436).
CC The sequences differ at nucleotides 619: T to G transversion. This clone
CC is not complete as the sequence starts 16 nucleotides downstream of the
CC initiator codon. The complete sequence has (AAT62440) has 18 exons, is
CC located at the D3F1S2 locus on human chromosome 3 and encodes for a
CC human growth factor which is an 80 kD single-chain protein containing 4
CC kringle domains. The protein can be used to alter cell growth (as a
CC growth factor or tumour suppressor) and has similar properties to the

CC hepatocyte growth factor that is actively involved in liver
CC regeneration.
XX
SQ Sequence 2219 BP; 461 A; 660 C; 674 G; 420 T; 4 other;
Query Match 47.9%; Score 810; DB 18; Length 2219;
Best Local Similarity 99.4%; Pred. No. 3.4e-191;
Matches 813; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 17 tctgtgcttctgactcaatgcttaggggtccctgggagcgctcgccattgaatgact 76
Db 1 tctgtgcttctgactcaataacttaggggtccctgggagcgctcgccattgaatgact 60
QY 77 tccaagtgtctccggggcacagagctacagcacctgctacatgcgtggtgctcccggtt 136
Db 61 tccaagtgtctccggggcacagagctacagcacctgctacatgcgtggtgctcccggtt 120
QY 137 ggcaggaggtggtggcagatgctgaagagtgtgctgctgctggtgggccccttaatggact 196
Db 121 ggcaggaggtggtggcagatgctgaagagtgtgctgctgctggtgggccccttaatggact 180
QY 197 gccgggcttccactacaacgtgagcagccatggttgccaaactgctccatggactcaac 256
Db 181 gccgggcttccactacaacgtgagcagccatggttgccaaactgctccatggactcaac 240
QY 257 actgccccacacagaggtgcggctgttctgggctgtgacctcttcagaagaagact 316
Db 241 actgccccacacagaggtgcggctgttctgggctgtgacctcttcagaagaagact 300
QY 317 acgtacggacctgcatcatgaacaatgggttgggtaccggggcaccatggccacgacg 376
Db 301 acgtacggacctgcatcatgaacaatgggttgggtaccggggcaccatggccacgacg 360
QY 377 tgggtggcctgccctgccaggcttgaggccacaagtctcccgaatgatcacaaagtacacg 436
Db 361 tgggtggcctgccctgccaggcttgaggccacaagtctcccgaatgatcacaaagtacacg 420
QY 437 ccactctccgggaatggcctggaagagaaacttctgcgttaacctgtggtgagacccccggag 496
Db 421 ccactctccgggaatggcctggaagagaaacttctgcgttaacctgtggtgagacccccggag 480
QY 497 gtccttgggtgtacacaaacagacctgtgtgctgttccagagctgcggcatcaaatcct 556
Db 481 gtccttgggtgtacacaaacagacctgtgtgctgttccagagctgcggcatcaaatcct 540
QY 557 gccgggagggccgctgtgtctggtgcaatggcgaggaataccgcggcggttagaccgca 616
Db 541 gccgggagggccgctgtgtctggtgcaatggcgaggaataccgcggcggttagaccgca 600
QY 617 cggagtcaggggcgcgagtgccagcgctgggatcttccagaccccgaccaccccttcg 676
Db 601 cggagtcaggggcgcgagttccagcgctgggatcttccagaccccgaccaccccttcg 660
QY 677 agccgggcaagtctcctcgaccaagtctggacgacaaactattgccggaatcctgacggct 736
Db 661 agccgggcaagtctcctcgaccaagtctggacgacaaactattgccggaatcctgacggct 720
QY 737 ccgagcgccatggtgtctacactacggtatccgcagatcgagcgagagttctgtgacctcc 796
Db 721 ccgagcgccatggtgtctacactacggtatccgcagatcgagcgagagttctgtgacctcc 780
QY 797 ccgctgcgggtccgagggcacagccccgcctcgagggc 834
Db 781 ccgctgcgggtccgagggcacagccccgcctcgaggcc 818

RESULT 12
AAT62438
ID AAT62438 standard; cDNA to mRNA; 2021 BP.
XX
AC AAT62438;
XX

PA (CHIL-) CHILDRENS HOSPITAL MEDICAL CENT.
XX
PI Degen SJ;
XX
DR WPI; 1994-166645/20.
XX P-PSDB; AAR66599.
XX
PT DNA from D3F15S2 locus of human chromosome 3 - encoding novel
PT growth factor, L5/3, useful as probe for detecting
PT pre-deposition towards cancer
XX
PS Disclosure; Columns 23-28; 31pp; English.
XX
CC The cDNA corresponding to a gene located at the D3F15S2 locus of
CC human chromosome 3, designated L5/3, has been isolated (see AAQ79723).
CC The gene codes for a protein composed of 4 kringle domains, followed
CC by a serine protease-like domain. The hepatocyte growth factor has a
CC similar structure but is only 50% identical to L5/3 at the amino
CC acid sequence level. The L5/3 protein is useful for altering cell
CC growth (as a growth factor or tumour suppressor). The L5/3 gene is
CC also identical to the gene deleted from all human small lung cell
CC carcinomas. The cDNA sequence of one L5/3 gene clone (i.e. clone #19,
CC AAQ79725) had two parts of the coding region deleted when compared to
CC the longest sequence AAQ79723. The region deleted included exon 13 and
CC the 5' end of exon 18.
XX
SQ Sequence 2021 BP; 426 A; 610 C; 612 G; 373 T; 0 other;

Query Match 46.8%; Score 792.6; DB 15; Length 2021;
Best Local Similarity 99.5%; Pred. No. 6.8e-187;
Matches 795; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 36 atgcttaggggtccctggcgagcgtcgccattggaatgaattccaagtctccggggcac 95
|||||
Db 1 atgcttaggggtccctggcgagcgtcgccattggaatgaattccaagtctccggggcac 60

QY 96 agagctacagcacctgctacatgcggtggtgccccggccttgccaggaggtgtggcaga 155
|||||
Db 61 agagctacagcacctgctacatgcggtggtgccccggccttgccaggaggtgtggcaga 120

QY 156 tgctgaagagtgctgctggtgctgtggcccttaattggaactgcgggccccttcactacaa 215
|||||
Db 121 tgctgaagagtgctgctggtgctgtggcccttaattggaactgcgggccccttcactacaa 180

QY 216 cgtgagcagccatggttgccaactgctgccatggactcaacactgcgccacacagagct 275
|||||
Db 181 cgtgagcagccatggttgccaactgctgccatggactcaacactgcgccacacagagct 240

QY 276 gcgcgcttctggcgctgtgacctcttcacagaagaagactacgtacggacctgcatcat 335
|||||
Db 241 gcgcgcttctggcgctgtgacctcttcacagaagaagactacgtacggacctgcatcat 300

QY 336 gaacaatggggttggttacccggggccaccatggccacacccgtgggtggcctgccctgcc 395
|||||
Db 301 gaacaatggggttggttacccggggccaccatggccacacccgtgggtggcctgccctgcc 360

QY 396 ggcttgagccacaagtcccgaaatgatcacaaagtacacgcccactctccggaatggcct 455
|||||
Db 361 ggcttgagccacaagtcccgaaatgatcacaaagtacacgcccactctccggaatggcct 420

QY 456 ggaagagaacttctgccgtaaccctgatggcgacccccggaggtccttggtgctacacaa 515
|||||
Db 421 ggaagagaacttctgccgtaaccctgatggcgacccccggaggtccttggtgctacacaa 480

QY 516 agacctgctgtgcgcttcagagctgcggcatcaaatcctgcgggagggcgcggtgtgt 575
|||||
Db 481 agacctgctgtgcgcttcagagctgcggcatcaaatcctgcgggagggcgcggtgtgt 540

QY 576 ctggtgcaatggcgaggaataaccgcggcggttagaccgacggagtcagggcgcgagtg 635
|||||
Db 541 ctggtgcaatggcgaggaataaccgcggcggttagaccgacggagtcagggcgcgagtg 600

QY 636 ccagcgtgggatcttcagcaccgcgaccagcacccttcgagccgggcaagtctcctcga 695
|||||
Db 601 ccagcgtgggatcttcagcaccgcgaccagcacccttcgagccgggcaagtctcctcga 660

QY 696 ccaaggtctggacgacaaactattgccgggaatcctgacggctccgagcggccatggtgcta 755
|||||
Db 661 ccaaggtctggacgacaaactattgccgggaatcctgacggctccgagcggccatggtgcta 720

QY 756 cactacggtatccgcagatcgagcgcgagagttctgtgacctccccgcgtgcgggtccgaggc 815
|||||
Db 721 cactacggtatccgcagatcgagcgcgagagttctgtgacctccccgcgtgcgggtccgaggc 780

QY 816 acagcccgctcgcagggc 834
|||||
Db 781 acagcccgcaagagggc 799

RESULT 14
AAx87676
ID AAX87676 standard; cDNA; 2190 BP.
XX
AC AAX87676;
XX
DT 26-OCT-1999 (first entry)
XX
DE HGF-MSP hybrid protein alphabet-1 factor cDNA.
XX
KW Alphabet-1; human; hepatocyte growth factor; HGF; MSP;
KW macrophage stimulating protein; apoptosis; chemotherapy; toxicity;
KW therapy; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key
FT sig_peptide 1..63
FT mat_peptide 63..2190
FT /*tag= a
FT /*tag= b
XX
PN WO9938968-A1.
XX
PD 05-AUG-1999.
XX
PF 28-JAN-1999; 99WO-EP00502.
XX
PR 30-JAN-1998; 98IT-MI00180.
XX
PA (DOMP-) DOMPE SPA.
XX
PI Caselli G, Collesi C, Comoglio P, Medico E, Michieli P;
XX
DR WPI; 1999-494091/41.
DR P-PSDB; AAY06621.
XX
PT Recombinant proteins from recombination of HGF and MSP structural
PT domains, useful for protecting cells against apoptosis induced by
PT chemotherapeutics
XX
PS Example 1; Page 52-53; 63pp; English.
XX
CC This DNA sequence codes for alphabet-1 factor (see AAY06621), a
CC recombinant protein composed of the hepatocyte growth factor (HGF)
CC alpha chain (i.e. the signal sequence, hairpin loop and kringles
CC 1-4), the natural cleavage site of HGF, the macrophage stimulating
CC protein (MSP) beta chain and a poly-histidine tag sequence. This
CC structure allows the recombinant protein to interact with both the
CC HGF receptor (Met) and the MSP receptor (Ron) and thereby induce
CC biological responses which are synergistic and selective compared
CC with the natural factor and truncated forms of the proteins. The
CC portions of DNA encoding the various portions of alphabet-1 are
CC obtained by PCR amplification of HGF or MSP cDNA and then
CC recombined to obtain the hybrid sequence. Expression vectors,

proliferation of the haematopoietic precursors or to stimulate their entry in the cycle. By modification of the proteolytic site, hybrid factors can be obtained which are activated by proteases of the endoplasmic reticulum (such as furines) during their synthesis. When the proteolytic site is removed, permanently immature forms of the factors can be obtained, having a potential partial agonistic or antagonistic activity. Different functional domains can be combined so as to modulate the biological effects.									
XX									
SQ	Sequence 2190 BP; 612 A; 528 C; 544 G; 506 T; 0 other;								
Query Match 45.7%; Score 773; DB 20; Length 2190; Best Local Similarity 98.7%; Pred. No. 5.1e-182; Matches 779; Conservative 0; Mismatches 10; Indels 0; Gaps 0;									
QY	874	tctctagagggacaaaaggaaaagaagaaatacaattcatgaattcaaaaaaatcagcaaag	933						
Db	82	tatgcagagggacaaaaggaaaagaagaaatacaattcatgaattcaaaaaaatcagcaaag	141						
QY	934	actaccctaatacaaaatagatccagcactgaagataaaaaaaccaaaaagtgaatactgca	993						
Db	142	actaccctaatacaaaatagatccagcactgaagataaaaaaaccaaaaagtgaatactgca	201						
QY	994	gaccaatgtgctaataatagatgtactaggaataaaaggacttccattgcagggtttt	1053						
Db	202	gaccaatgtgctaataatagatgtactaggaataaaaggacttccattgcagggtttt	261						
QY	1054	gtttttgataaaagcaagaaaacaatgcctctgtgttccccttcaatagcatgtcaagtga	1113						
Db	262	gtttttgataaaagcaagaaaacaatgcctctgtgttccccttcaatagcatgtcaagtga	321						
QY	1114	gtgaaaaagaatttgccatgaatttgacctctatgaaaaacaaagactacattagaaac	1173						
Db	322	gtgaaaaagaatttgccatgaatttgacctctatgaaaaacaaagactacattagaaac	381						
QY	1174	tgcattcattggttaaaggacgcagctacaagggaacagtatctatcactaagagtggcatc	1233						
Db	382	tgcattcattggttaaaggacgcagctacaagggaacagtatctatcactaagagtggcatc	441						
QY	1234	aaatgtcagccctggaggttccatgatataccacgaacagctatcggggtaaagacct	1293						
Db	442	aaatgtcagccctggaggttccatgatataccacgaacagctatcggggtaaagacct	501						
QY	1294	caggaaaaactactgtcgaaatcctcagagggaagaaaggggacccctggtgtttcacaagc	1353						
Db	502	caggaaaaactactgtcgaaatcctcagagggaagaaaggggacccctggtgtttcacaagc	561						
QY	1354	aatccagaggttacgcctacgaagtctgtgacattcctcagttcagaagttgaatgc	1413						
Db	562	aatccagaggttacgcctacgaagtctgtgacattcctcagttcagaagttgaatgc	621						
QY	1414	acctgcaatggggagaggttatcgagggtctcatggatcatacagaatacaggcaagattt	1473						
Db	622	acctgcaatggggagaggttatcgagggtctcatggatcatacagaatacaggcaagattt	681						
QY	1474	cagcgctgggatcatcagacacacacacccggcacaaattcttgctgaaagatatcccgac	1533						
Db	682	cagcgctgggatcatcagacacacacacccggcacaaattcttgctgaaagatatcccgac	741						
QY	1534	aagggttttgatgataaattattgcccgaatccccgatggccagccgagggccatggtgct	1593						
Db	742	aagggttttgatgataaattattgcccgaatccccgatggccagccgagggccatggtgct	801						
QY	1594	actcttgacccctcacaccccgctgggagtagtctgtgcaattaaaaacatgcgctgacaaa	1653						
Db	802	actcttgacccctcacaccccgctgggagtagtctgtgcaattaaaaacatgcgctgacaaa	861						
QY	1654	gacgacgac	1662						
Db	862	atgaatgac	870						

Search completed: September 18, 2001, 18:34:03
Job time: 7520 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 18:29:31 ; Search time 97.89 Seconds
(without alignments)
3272.187 Million cell updates/sec

Title: US-09-600-991-3
Perfect score: 1692
Sequence: 1 atggggtggctccactcct.....accaccaccaccactag 1692

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCtUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	829.2	49.0	2232	1 US-08-334-177-1	Sequence 1, Appli
2	829.2	49.0	2232	5 PCT-US95-13830-1	Sequence 1, Appli
3	827.6	48.9	2216	2 US-08-666-082B-2	Sequence 2, Appli
4	811.6	48.0	2219	1 US-07-882-925A-1	Sequence 1, Appli
5	811.6	48.0	2219	1 US-08-184-012C-1	Sequence 1, Appli
6	811.6	48.0	2262	1 US-07-882-925A-7	Sequence 7, Appli
7	811.6	48.0	2262	1 US-08-184-012C-7	Sequence 7, Appli
8	810	47.9	2219	1 US-07-882-925A-2	Sequence 2, Appli
9	810	47.9	2219	1 US-08-184-012C-2	Sequence 2, Appli
10	794.2	46.9	2021	1 US-07-882-925A-3	Sequence 3, Appli
11	794.2	46.9	2021	1 US-08-184-012C-3	Sequence 3, Appli
12	773	45.7	2289	1 US-07-838-410-2	Sequence 2, Appli
13	771.4	45.6	2172	4 US-08-030-410-2	Sequence 2, Appli
14	761	45.0	2288	1 US-08-290-937B-4	Sequence 4, Appli
15	748	44.2	2184	1 US-07-815-333A-1	Sequence 1, Appli
16	533	31.5	2188	1 US-07-882-925A-4	Sequence 4, Appli
17	533	31.5	2188	1 US-08-184-012C-4	Sequence 4, Appli
18	284.8	16.8	6100	1 US-07-882-925A-6	Sequence 6, Appli
19	284.8	16.8	6100	1 US-08-184-012C-6	Sequence 6, Appli
20	136.4	8.1	1284	3 US-08-985-526-24	Sequence 24, Appli
21	136.2	8.0	645	3 US-08-985-526-22	Sequence 22, Appli
22	136	8.0	2497	1 US-08-643-219-12	Sequence 12, Appli
23	136	8.0	2497	2 US-09-131-995-12	Sequence 12, Appli
24	136	8.0	2497	2 US-08-832-087B-12	Sequence 12, Appli
25	136	8.0	2497	3 US-08-851-350-12	Sequence 12, Appli
26	136	8.0	2497	4 US-09-132-154-12	Sequence 12, Appli
27	136	8.0	2679	6 5200340-7	Patent No. 5200340

28	134.4	7.9	2753	1	US-07-854-603-1	Sequence 1, Appli
29	134	7.9	1134	4	US-09-206-059-29	Sequence 29, Appli
30	128.4	7.6	2296	1	US-07-750-080A-18	Sequence 18, Appli
31	128.4	7.6	2296	3	US-08-651-472-18	Sequence 18, Appli
32	114	6.7	6751	1	US-07-882-925A-5	Sequence 5, Appli
33	114	6.7	6751	1	US-08-184-012C-5	Sequence 5, Appli
34	60.4	3.6	1869	3	US-08-952-967-7	Sequence 7, Appli
35	60.2	3.6	4092	2	US-08-469-537A-106	Sequence 106, App
36	58	3.4	1947	1	US-07-998-972A-2	Sequence 2, Appli
37	58	3.4	1947	1	US-08-463-953-2	Sequence 2, Appli
38	58	3.4	1947	1	US-08-462-261-2	Sequence 2, Appli
39	58	3.4	1947	2	US-08-479-733A-24	Sequence 24, Appli
40	58	3.4	1947	3	US-08-487-427-24	Sequence 24, Appli
41	58	3.4	1947	3	US-08-479-727A-24	Sequence 24, Appli
42	58	3.4	1947	3	US-08-482-369A-24	Sequence 24, Appli
43	58	3.4	1947	5	PCT-US92-11357-2	Sequence 24, Appli
44	58	3.4	1947	5	PCT-US95-07439-24	Sequence 24, Appli
45	58	3.4	1988	1	US-07-750-080A-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1
US-08-334-177-1
; Sequence 1, Application US/08334177
; Patent No. 5696086
; GENERAL INFORMATION:
; APPLICANT: Avraham, Hava Karsenty
; APPLICANT: Godowski, Paul J.
; TITLE OF INVENTION: Methods and Kits Using Macrophage Stimulating Protein
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: . PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/334,177
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: 912
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-5416
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2232 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-334-177-1

Query Match 49.0%; Score 829.2; DB 1; Length 2232;
Best Local Similarity 99.6%; Pred. No. 2.6e-211;
Matches 831; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 atggggtggctccactcctgctgtcttctgactcaatgtaggggtccctgggcagcgc 60

QY 661 caccagacccttcgagccgggcaagtctcctcgaccaaggtctggtgacgacaaactattgc 720
|||||
Db 671 CACCAGACCCCTTCGAGCCGGCAAGTTCTCGACCAAGGTCTGGACGACAACTATTGC 730
QY 721 cggaatcctgacggtccgagcggccatggtgtctacactacggtatccgcagatcgagcga 780
|||||
Db 731 CGGAATCCTGACGGCTCCGAGCGGCATGGTGCTACACTACGGATCCGCAGATCGAGCGA 790
QY 781 gagtctgtgacctcccccgctcggggtccgagggcacagcccccgctcgagggc 834
|||||
Db 791 GAGTTCTGTGACCTCCCCCGCTGGGGTCCGAGGCACAGCCCCCGCCAAAGAGGCC 844

RESULT 3

US-08-666-082B-2
; Sequence 2, Application US/08666082B
; Patent No. 5916770
; GENERAL INFORMATION:
; APPLICANT: YOSHIKAWA, WATARU
; APPLICANT: SHIMONISHI, MANABU
; APPLICANT: IWAMOTO, JUNKO
; APPLICANT: TAKEHARA, TOYOHIRO
; APPLICANT: HAGIYA, MICHIO
; TITLE OF INVENTION: MACROPHAGE STIMULATING PROTEIN VARIANT AND
; TITLE OF INVENTION: METHOD FOR PRODUCING THE SAME
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LEYDIG, VOIT & MAYER, LTD.
; STREET: TWO PRUDENTIAL PLAZA, SUITE 4900
; CITY: CHICAGO
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/666,082B
; FILING DATE: 19-JUN-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-153309
; FILING DATE: 20-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBERT F. GREEN
; REGISTRATION NUMBER: 27555
; REFERENCE/DOCKET NUMBER: 73843
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2216 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Human
; FEATURE:
; NAME/KEY: sig peptide
; LOCATION: 1..54
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2133
US-08-666-082B-2

Query Match 48.9%; Score 827.6; DB 2; Length 2216;
Best Local Similarity 99.5%; Pred. No. 6.8e-211;
Matches 830; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 atgggggggtccctcctgctgtctgtgactcaatgcttaggggtccctgggcagcgc 60
|||||

Db 1 ATGGGGTGGTCCCACTCTCTGCTGCTTCTGACTCAATACTTAGGGGTCCCTGGGCGAGCGC 60
QY 61 tcgccattgaatgacttccaagtctccggggtcacagagctacagcacctgctacatgcg 120
|||||
Db 61 TCGCCATTGAATGACTTCCAAGTGCTCCGGGGCACAGAGTACAGCACCTGCTACATGCG 120
QY 121 gtggtgccccggccttgccaggagatgtggcagatgctgaagagtgtgctggtcgctgt 180
|||||
Db 121 GTGGTCCCCGGGCTTGGCAGGAGGATGTGGCAGATGCTGAAGAGTGTGCTGGTGGCTGT 180
QY 181 gggcccttaatgactgcccgggccccttcactacaacgtgagcagccatggttgccaaactg 240
|||||
Db 181 GGGCCCTTAATGGACTGCCGGGCCCTTCCACTACAACGTGAGCAGCCATGGTTGCCAACTG 240
QY 241 ctgccatggactcaacactcgcgccacacagaggtcggttcttgggctgtgacctc 300
|||||
Db 241 CTGCCATGGACTCAACACTCGCCCCCACACGAGGCTCGGGCGTCTTGGGCGCTGTGACCTC 300
QY 301 ttccagaagaagactacgtacggacctgcatcatgaacaatgggttgggtaccggggc 360
|||||
Db 301 TTCCAGAAGAAGACTACGTACGGACCTGCATCATGAACAATGGGTTGGGTACCCGGGCG 360
QY 361 accatggccacgacccgtgggtggcctgcctgccctgccaggttggagccacaagtcccgaa 420
|||||
Db 361 ACCATGGCCACGACCGTGGTGGCTGCCCTGCCAGGCTTGGAGCCACAAGTTCCCGAAT 420
QY 421 gatcacaagtacacgcccactctccggaatggcctggaagagaacttctgcgtaacct 480
|||||
Db 421 GATCACAAGTACACGCCCACTCTCCGGAATGGCTGGAAGAGAACTTCTGCCGTAAACCT 480
QY 481 gatggcagccccggaggtccttgggtctacacaaacagacctgctgtgcgttccagagc 540
|||||
Db 481 GATGGCGACCCCGGAGTCTTGGTGCTACACAACAGACCTGCTGTGCGCTTCCAGAGC 540
QY 541 tgcggcatcaaatcctgcccgggagggcgctgtgtctgtgcaatggcgaggaataccgc 600
|||||
Db 541 TCGGGCATCAAAATCCTGCGGGAGGCCGCGTGTGTCTGTGCAATGGCGAGGAATACCGC 600
QY 601 ggcggttagaccgcaaggagtcaggcgagtcagggcgagtcgccagcgtgggtatcttcagcacccg 660
|||||
Db 601 GCGCGGTAGACCCGACGGAGTCAAGGCGCGAGTGCCAGCGCTGGGATCTTCAGCACCCG 660
QY 661 caccagaccccttcgagccgggcaagtctcctcgaccaagtctggtgacgacaaactattgc 720
|||||
Db 661 CACCAGACCCCTTCGAGCCGGGCAAGTTCTCTCGACCAAGTCTGGACGACAACTATTGC 720
QY 721 cggaatcctgacggtcctcgagcggccatggtgtctacactacggtatccgcagatcgagcga 780
|||||
Db 721 CGGAATCCTGACGGTCCGAGCGGCCCATGGTGCTACACTACGGATCCGCAGATCGAGCGA 780
QY 781 gagtctgtgacctcccccgctcggggtccgagggcacagcccccgctcgagggc 834
|||||
Db 781 GAGTTCTGTGACCTCCCCCGCTGGGGTCCGAGGCACAGCCCCCGCCAAAGAGGCC 834

RESULT 4

US-07-882-925A-1
; Sequence 1, Application US/07882925A
; Patent No. 5315000
; GENERAL INFORMATION:
; APPLICANT: Degen, Sandra J. F.
; TITLE OF INVENTION: Gene for a growth factor and its cDNA and
; TITLE OF INVENTION: protein
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gregory Lunn
; STREET: Wood, Herron & Evans, 2700 Carew Tower
; CITY: Cincinnati
; STATE: Ohio
; COUNTRY: USA
; ZIP: 45202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb

QY	617	cggagtcagggcgcgagtgccagcgctgggatcttcagcaccgcaccagcacccttcg	676
Db	632	CGGAGTCAGGGCGCGAGTGCCAGCGCTGGGATCTTCAGCACCCGACACGACCCCTTCG	691
QY	677	agccgggcaagttcctcgaccaaaggtctggacgacaactattgcgggaatcctgacggct	736
Db	692	AGCCGGGCAAGTTCCTCGACCAAGGTCTGGACGACAACATAATGCCGGAATCCTGACGGCT	751
QY	737	ccgagcggccatggtgctacactacggatccgcagatcgagcgagaggttctgtgacctcc	796
Db	752	CCGAGCGGCCATGGTGTCTACACTACGGATCCGCAGATCGAGCGAGAGTTCTGTGACCTCC	811
QY	797	cccgtcggggtccgagggcacagccccgcctcgagggc	834
Db	812	CCCGCTGCGGGTCCGAGGCGACAGCCCCGCCCAAGAGGCC	849

RESULT 8
 US-07-882-925A-2
 ; Sequence 2, Application US/07882925A
 ; Patent No. 5315000
 ; GENERAL INFORMATION:
 ; APPLICANT: Degen, Sandra J. F.
 ; TITLE OF INVENTION: Gene for a growth factor and its cDNA and
 ; TITLE OF INVENTION: protein
 ; NUMBER OF SEQUENCES: 7
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Gregory Lunn
 ; STREET: Wood, Herron & Evans, 2700 Carew Tower
 ; CITY: Cincinnati
 ; STATE: Ohio
 ; COUNTRY: USA
 ; ZIP: 45202
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb
 ; COMPUTER: Apple Macintosh
 ; OPERATING SYSTEM: Macintosh 6.0.3
 ; SOFTWARE: Microsoft Word 4.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07882,925A
 ; FILING DATE: 19920514
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lunn, Gregory
 ; REGISTRATION NUMBER: 29,945
 ; REFERENCE/DOCKET NUMBER: CMC 57
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (513) 241-2324
 ; TELEFAX: (513) 421-7269
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2219 base pairs
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; ANTI-SENSE: no
 ; ORIGINAL SOURCE:
 ; ORGANISM: human
 ; DEVELOPMENTAL STAGE: fetal
 ; TISSUE TYPE: liver
 ; IMMEDIATE SOURCE:
 ; LIBRARY: cDNA
 ; CLONE: #33
 ; POSITION IN GENOME:
 ; CHROMOSOME/SEGMENT: human 3p21/D3F15S2
 ; FEATURE:
 ; IDENTIFICATION METHOD: experimental
 ; OTHER INFORMATION: Includes five polymorphisms at the nucleotide
 ; OTHER INFORMATION: level; one of which results in an amino acid substitution (nu
 ; OTHER INFORMATION: 619). Sequence ID NO:1: contains the identical sequence with
 ; OTHER INFORMATION: polymorphic amino acid.
 ; PUBLICATION INFORMATION:

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2021 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; DEVELOPMENTAL STAGE: fetal
; TISSUE TYPE: liver
; IMMEDIATE SOURCE:
; LIBRARY: cDNA
; CLONE: #19
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: human 3p21/D3F15S2
; FEATURE:
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: This sequence is a variant where two regions
; OTHER INFORMATION: were found to be deleted when compared to SEQ ID NO:1.
; PUBLICATION INFORMATION:
; RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 2021
US-07-882-925A-3

Query Match 46.9%; Score 794.2; DB 1; Length 2021;
Best Local Similarity 99.6%; Pred. No. 5.1e-202;
Matches 796; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 36 atgcttaggggtccctggcagcgctgcgcaattgaatgacttccaaagtctccggggcac 95
Db 1 ATGCTTAGGGGTCCCTGGCAGCGCTCGGCATTTGAATGACTTCCAAGTCTCCGGGGCAC 60

QY 96 agagctacagcacctgctacatgcggtggtgcccgggccccttggcaggaggtgtggcaga 155
Db 61 AGAGCTACAGCACCTGCTACATGCGGTGTCGCCGGCCCTTGGCAGGAGGATGTGGCAGA 120

QY 156 tgctgaagagtgtgctggtcgctgtgggccccttaatggactgcgggccccttcactacaa 215
Db 121 TGCTGAAGAGTGTGCTGGTCGCTGTGGGCCCTTAATGGACTGCCGGGCTTCCACTACAA 180

QY 216 cgtgagcagccatggttgccaaactgctgccatggactgaactcaacactgcgccacacaggct 275
Db 181 CGTGAGCAGCCATGGTTGCCAACTGCTGCCATGGACTCAACACTCGCCACACAGAGGCT 240

QY 276 gcggcgttctggcgctgtgacctcttcagaaagaaactacgtacggacctgcatcat 335
Db 241 CGGGCGTTCTGGCGCTGTGACCTCTTCAGAAAGAAAGACTACGTACGGACCTGCATCAT 300

QY 336 gaacaatggggttggtaccgggggaccatggccacgacctgggtggcctgccctgccca 395
Db 301 GAACAATGGGGTGGGTACCGGGGCACCATGGCCACGACCGTGGGTGGCTGCCCTGCCA 360

QY 396 ggcttgagccacaagtctccgaatgatacaagaatcacacgcccactctcgggaatggcct 455
Db 361 GGCTTGAGCCACAAGTTCCCGAATGATCACAAAGTACACGCCCACTCTCCGGGAATGGCCT 420

QY 456 ggaagagaacttctgccgttaaccctgatggcagaccccgagggtccttgctgctacacaac 515
Db 421 GGAAGAGAACTTCTGCGGTAACCCCTGATGGCGGACCCCGGAGGTCTCTGTGTGCTACACAAC 480

QY 516 agaccctgctgctgcttccagagctgcggcgcatacaaatcctgcgggagccgcgtgtgt 575
Db 481 AGACCCCTGCTGTGGCTTCCAGAGCTGCGGCATCAAAATCCTGCCGGGAGSCCGGTGTGT 540

QY 576 ctggtgcaatggcgaggaataaccgcggcggttagaccgcacggagtgtagggcgcgagtgtg 635
Db 541 CTGGTGAATGGCGAGGAATACCGCGGGCGCGGTAGACCCGACGGAGTGTCAGGGCGCGAGTG 600

QY 636 ccagcgctgggatcttcagcaccgcgacccgaccccttcgagccggcggaagtctcctcga 695
Db 601 CCAGCGCTGGGATCTTCAGCACCCGACCCAGCACCCCTTCGAGCCGGGCAAGTTCCTCTCGA 660

QY 696 ccaaggtctggacgacaaactattgcccgaatacctgacggctccgagcgccatggtgcta 755
Db 661 CCAAGGTCTGGACGACAACTATTGCCGGAATCCTGACGGCTCCGAGCGGCCATGGTGCTA 720

QY 756 cactacggatccgcagatcgagcgagaggttctgtgacctcccccgctgcgggtccgaggc 815
Db 721 CACTACGGATCCGCGAGATCGAGCGAGAGTCTGTGACCTCCCGCGTCCGCGGTCCGAGGC 780

QY 816 acagccccgcctcgagggc 834
Db 781 ACAGCCCCGCCAAGAGGCC 799

RESULT 11
US-08-184-012C-3
; Sequence 3, Application US/08184012C
; Patent No. 5606029
; GENERAL INFORMATION:
; APPLICANT: Degen, Sandra J. F.
; TITLE OF INVENTION: Gene for a growth factor and its cDNA and
; TITLE OF INVENTION: protein
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gregory Lunn
; STREET: Wood, Herron & Evans, 2700 Carew Tower
; CITY: Cincinnati
; STATE: Ohio
; COUNTRY: USA
; ZIP: 45202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 800 Kb
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5.2
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/184,012C
; FILING DATE: 1/18/94
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Gregory
; REGISTRATION NUMBER: 29,945
; REFERENCE/DOCKET NUMBER: CMC 57
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (513) 241-2324
; TELEFAX: (513) 421-7269
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2021 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; ANTI-SENSE: no
; ORIGINAL SOURCE:
; ORGANISM: human
; DEVELOPMENTAL STAGE: fetal
; TISSUE TYPE: liver
; IMMEDIATE SOURCE:
; LIBRARY: CDNA
; CLONE: #19
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT: human 3p21/D3F15S2
; FEATURE:
; IDENTIFICATION METHOD: experimental
; OTHER INFORMATION: This sequence is a variant where two regions
; OTHER INFORMATION: were found to be deleted when compared to SEQ ID NO:1.
; PUBLICATION INFORMATION:
; RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 2021
US-08-184-012C-3

Query Match 46.9%; Score 794.2; DB 1; Length 2021;

```
QY 1234 aaatgtcagccctggagttccatgataccacagaacacagctatcggggtaaaagaccta 1293
|||||
Db 519 AAATGTCAGCCCTGGAGTTCCATGATACCACAGAACACAGCTATCGGGTAAAGACCTA 578

QY 1294 caggaanaactactgtcgaaatcctcgaggggaagaaggggacccctggtgttcacaagc 1353
|||||
Db 579 CAGGAAAACACTGTGCGAAATCCTCGAGGGGAAGAAGGGGACCCCTGGTGTTCACAAAGC 638

QY 1354 aatccagaggtacgtctacgaagtctgtgacattcctcagtgttcagaagttgaatgcatg 1413
|||||
Db 639 AATCCAGAGGTACGCTACGAAGTCTGTGACATTCTCAGTGTTCAGAAAGTTGAATGCATG 698

QY 1414 acctgcaatggggagagttatcgaggtctctcatggatcatatcacagaatcaggcgaagatttgt 1473
|||||
Db 699 ACCTGCAATGGGAGAGTTATCGAGGTCTCATGGATCATACAGAAATCAGCAAGATTGTGT 758

QY 1474 cagcgtgggatcatcatcagacaccacacccggcacaaaattcttgccctgaaagatatcccgac 1533
|||||
Db 759 CAGCGCTGGGATCATCAGACACCACACCCGGCACAAAATTCTTGCCCTGAAAGATATCCCGAC 818

QY 1534 aagggctttgatgataattattgcccgaatcccgatggccagccgagggccatggtgctat 1593
|||||
Db 819 AAGGGCTTTGATGATAATTATTGCCGCAATCCCGATGCCAGCCGAGGCCATGGTGCTAT 878

QY 1594 actcttgaccctcacaccccgctgggagtgactgtgcaattaaacaatcgctgacaaagct 1653
|||||
Db 879 ACTCTTGACCCTCACACCCCGCTGGGAGTACTGTGCAATTAAACATCGCGTGACAATACT 938

QY 1654 gacgacgac 1662
|||||
Db 939 ATGAATGAC 947
```

RESULT 13

```
US-08-030-410-2
; Sequence 2, Application US/08030410
; Patent No. 6221359
; GENERAL INFORMATION:
; APPLICANT: Komiya, Atsushi
; APPLICANT: Nakahata, Tatsutoshi
; APPLICANT: Kubo, Tetsuo
; APPLICANT: Tanaka, Ryuhei
; APPLICANT: Kawano, Genji
; APPLICANT: Sudo, Tetsuo
; APPLICANT: Sano, Emiko
; APPLICANT: Kojima, Katsuaki
; TITLE OF INVENTION: HEMOPOIETIC STEM CELL MULTIPLIER
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/030,410
; FILING DATE: 19930521
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8898
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
```

; INFORMATION FOR SEQ ID NO: 2:

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2172 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2169
US-08-030-410-2
```

```
Query Match 45.6%; Score 771.4; DB 4; Length 2172;
Best Local Similarity 98.6%; Pred. No. 6.2e-196;
Matches 778; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 874 tctctagagggacaaaaggaaaagaagaataacaaattcatgaattcaaaaaatcagcaaaag 933
|||||
Db 82 TATGCAGAGGGACAAAAGGAAAAGAAATAACAATTTCATGAATTCAAAAAATCAGCAAAG 141

QY 934 actaccctaatacaaaatagatccagcactgaagataaaaaacccaaaaaagtgaatactgca 993
|||||
Db 142 ACTACCCTAATCAAAATAGATCCAGCACTGAAGATAAAAAACCAAAAAAGTGAATACTGCA 201

QY 994 gaccaatgtctaataatagatgtactaggaaataaaaggacttccattcacttgcaaggctttt 1053
|||||
Db 202 GACCAATGTCTAATAGATGTACTAGGAATAAAGGACTTCCATTCTACTTCAAGGCTTTT 261

QY 1054 gtttttgataaaagcaagaaaaaataatgcctctctggttcccttcaatagcatgtcaagtgg 1113
|||||
Db 262 GTTTTGTGATAAAGCAAGAAAACAATGCTCTGGTTCCCTTCAATAGCATGTCAAGTGA 321

QY 1114 gtgaaaaaagaatttggccatgaatttgacctctatgaaaaacaaagactacattagaaac 1173
|||||
Db 322 GTGAAAAAAGAATTTGGCCATGAATTTGACCTCTATGAAAAACAAGACTACATTAGAAAC 381

QY 1174 tgcattcattggtaaaggacgcagctacaaagggaacagtatctatcactaaagtggtgcatc 1233
|||||
Db 382 TGCATCATTTGTTAAAGGACGCGAGCTACAAGGGAACAGTATCTATCACTAAGAGTGGCATC 441

QY 1234 aaatgtcagccctggagttccatgataccacacgaacacagctatcggggtaaagacct 1293
|||||
Db 442 AAATGTCAGCCCTGGAGTTCCATGATACCACACGAACACAGCTATCGGGSTAAAGACCTA 501

QY 1294 caggaaaaactactgtcgaaaatcctcgagggggaagaaggggacccctggtgttcacaagc 1353
|||||
Db 502 CAGGAAAACACTACTGTGCGAAATCCTCGAGGGGAAGAAGGGGACCCCTGGTGTTCACAAGC 561

QY 1354 aatccagaggtacgtacgaagtctgtgacattcctcagtgttcagaagttgaatgcatg 1413
|||||
Db 562 AATCCAGAGGTACGCTACGAAGTCTGTGACATTCTCAGTGTTCAGAAAGTTGAATGCATG 621

QY 1414 acctgcaatggggagagttatcgaggtctctcatggatcatatcacagaatcaggcgaagattgt 1473
|||||
Db 622 ACCTGCAATGGGAGAGTTATCGAGGTCTCATGGATCATACAGAAATCAGCAGGATTGTGT 681

QY 1474 cagcgtgggatcatcatcagacaccacacccggcacaaaattcttgccctgaaagatatcccgac 1533
|||||
Db 682 CAGCGCTGGGATCATCAGACACCACACCCGGCACAAATTCCTGCTGAAAGATATCCCGAC 741

QY 1534 aagggctttgatgataattattgcccgaatcccgatggccagccgagggccatggtgctat 1593
|||||
Db 742 AAGGGCTTTGATGATAATTATTGCCGCAATCCCGATGCCAGCCGAGGCCATGGTGCTAT 801

QY 1594 actcttgaccctcacaccccgctgggagtgactgtgcaattaaacaatcgctgacaaagct 1653
|||||
Db 802 ACTCTTGACCCTCACACCCCGCTGGGAGTACTGTGCAATTAAACATCGCGTGACAATACT 861

QY 1654 gacgacgac 1662
|||||
Db 862 ATGAATGAC 870
```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 2184 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2184
; US-07-815-333A-1

Search completed: September 18, 2001, 18:30:08
Job time: 7701 sec

Query Match		44.2%;	Score 748;	DB 1;	Length 2184;
Best Local Similarity		96.9%;	Pred. NO. 1.1e-189;		
Matches 779;		Conservative 0;	Mismatches 10;	Indels 15;	Gaps 1;
QY	874	tctctagagggaagaaagaaatacaattcatgaattcaaaaaatcagcaaaag	933		
Db	82	TATGCAGAGGGACAAAGGAAAGAAAGAAATACAATTTCATGAATTCAAAAATCAGCAAAAG	141		
QY	934	actaccctaatacaaaatagatccagcactgaagataaaaaaaccaaaaagtgaatactgca	993		
Db	142	ACTACCCCTAATCAAAATAGATCCAGCACTGAAGATAAAACCAAAAAGTGAATACTGCA	201		
QY	994	gaccaatgtgctaataagatgtactaggaataaaggacttccattccacttgcaaggctttt	1053		
Db	202	GACCAATGTGCTAATAGATGTACTAGGAATAAAGGACTTCCATTCACTTGCAAGGCTTTT	261		
QY	1054	gtttttgataaagcaagaaacaatgcctctgttcccccttcaatagcatgtcaagtgga	1113		
Db	262	GTTTTGTATAAAGCAAGAAAAACAATGCCTCTGTGTTCCCTTCAATAGCATGTCAAGTGGA	321		
QY	1114	gtgaaaaaagaatttggcccatgaatttgacctctatgaaaaacaaagactacattagaaaac	1173		
Db	322	GTGAAAAAAGAATTTGGCCATGAATTTGACCTCTATGAAAAACAAAGACTACATTAGAAAAC	381		
QY	1174	tgcatacttggtaaaggacgcagctacaaaggaacagtatctatcactaaagatggcatc	1233		
Db	382	TGCATCATTTGGTAAGGACGCAGCTACAAGGGACAGTATCTATCACTAAGAGTGGCATC	441		
QY	1234	aaatgtcagccctggagttccatgataccacacacgaacac-----agctat	1278		
Db	442	AAATGTCAGCCCTGGAGTTCCATGATACCACACGACACAGCTTTTTCGCTTCGAGCTAT	501		
QY	1279	cggggtaaagaccctacaggaataactactgtcgaatcctcaggggaagaaaggggaccc	1338		
Db	502	CGGGGTAAAGACCTACAGGAAAAAATACTGTGCAATCCTCGAGGGGAAGAGGGGACCC	561		
QY	1339	tgggtttcacaaagcaatccagaggtacgctacgaagtctgtgacattcctcagtgttca	1398		
Db	562	TGGTGTTTTCAACAGCAATCCAGAGGTACGCTACGAAGTCTGTGACATTCTCTCAGTGTCA	621		
QY	1399	gaagtgaatgcatgacctgcaatggggagagttatcgaggctctcatggatcacagaa	1458		
Db	622	GAAGTTGAATGCATGACCTGCAATGGGGAGAGTTATCGAGGTCTCATGGATCATACAGAA	681		
QY	1459	tcaggcaagatttgtcagcgctgggatcatcagacaccacaccggcacaaattcttgccct	1518		
Db	682	TCAGGCAAGATTTTGTACGGCTGGGATCATCAGACACCACACCGGCACAAATCTTGCCT	741		
QY	1519	gaaagatatcccgaaagggcttttgatgataattattgccgcaatcccgatggccagccg	1578		
Db	742	GAAAGATATCCCGACAAGGGCTTTTGATGATAATTTATGCCGAATCCCGATGGCCAGCCG	801		
QY	1579	aggccatggtgctatactcttgacctcacaccctgggagtgactgtgcaattaaaaaca	1638		
Db	802	AGGCCATGGTGTCTACTCTTGACCCCTCACACCCCTGGGAGTACTGTGCAATTAAAAACA	861		
QY	1639	tgcgctgacaaagctgacgacgac	1662		
Db	862	TGCGCTGACAATACTATGAATGAC	885		

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 18:27:22 ; Search time 2582.57 seconds
(without alignments)
6193.144 Million cell updates/sec

Title: US-09-600-991-3
Perfect score: 1692
Sequence: 1 atgggggtggtccactct.....accaccaccaccactag 1692

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- EST:*
- 1: gb_est1:*
 - 2: gb_est2:*
 - 3: gb_est3:*
 - 4: gb_est4:*
 - 5: gb_est5:*
 - 6: gb_est6:*
 - 7: gb_est7:*
 - 8: gb_est8:*
 - 9: gb_est9:*
 - 10: gb_est10:*
 - 11: gb_est11:*
 - 12: gb_est12:*
 - 13: gb_est13:*
 - 14: gb_est14:*
 - 15: gb_est15:*
 - 16: gb_est16:*
 - 17: gb_est17:*
 - 18: gb_est18:*
 - 19: gb_est19:*
 - 20: gb_est20:*
 - 21: gb_est21:*
 - 22: gb_est22:*
 - 23: gb_est23:*
 - 24: gb_est24:*
 - 25: gb_est33:*
 - 26: gb_est34:*
 - 27: gb_est35:*
 - 28: gb_est36:*
 - 29: gb_est37:*
 - 30: gb_est38:*
 - 31: gb_est39:*
 - 32: gb_est40:*
 - 33: em_estba:*
 - 34: em_estfun:*
 - 35: em_esthum1:*
 - 36: em_esthum2:*
 - 37: em_esthum3:*
 - 38: em_esthum4:*
 - 39: em_esthum5:*
 - 40: em_esthum6:*
 - 41: em_esthum7:*
 - 42: em_esthum8:*
 - 43: em_esthum9:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estrol:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estrol0:*
- 92: em_estrol1:*
- 93: em_estrol2:*
- 94: em_estrol3:*
- 95: em_estrol4:*
- 96: em_estrol5:*
- 97: em_estrol6:*
- 98: em_estrol7:*
- 99: em_estrol8:*
- 100: em_estrol9:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	736	43.5	975	106	AL546558	AL546558 AL546558
2	707	41.8	919	106	AL549164	AL549164 AL549164
3	594.4	35.1	872	106	AL550803	AL550803 AL550803
4	526.2	31.1	828	106	AL546883	AL546883 AL546883
5	475.2	28.1	766	148	BF383652	BF383652 602044540
6	456	27.0	889	106	AL558014	AL558014 AL558014
7	455.8	26.9	780	146	BF234001	BF234001 602024489
8	446	26.4	734	148	BF384148	BF384148 602045784
9	428.4	25.3	652	148	BF385836	BF385836 602046003
10	414.2	24.5	844	106	AL546856	AL546856 AL546856
11	405.4	24.0	543	139	BE723534	BE723534 193407 MA
12	389.2	23.0	899	106	AL573472	AL573472 AL573472
13	350.2	20.7	506	1	AA037738	AA037738 zf03b09.r
14	349.6	20.7	518	139	BE755132	BE755132 208867 MA
15	230.2	13.6	263	121	AW815873	AW815873 QV0-ST021
16	207.4	12.3	574	165	BE234370	BE234370 141225 MA
17	184.4	10.9	412	110	AV759905	AV759905 AV759905
18	175.6	10.4	288	161	BB598484	BB598484 BB598484
19	166	9.8	484	1	AA037786	AA037786 zf03b09.s
20	153.2	9.1	661	112	AW175268	AW175268 fi34b04.y
21	145.6	8.6	830	106	AL534044	AL534044 AL534044
22	135.8	8.0	753	227	AQ381428	AQ381428 RPC11-13
23	131.6	7.8	778	18	AI317783	AI317783 uj20f12.y
24	131.4	7.8	565	144	BF079866	BF079866 230570 MA
25	128	7.6	768	18	AI317745	AI317745 uj20b05.y
26	125.8	7.4	837	16	AI118965	AI118965 ue93b02.y
27	124.4	7.4	844	21	AI530175	AI530175 ui90a05.y
28	122.8	7.3	470	232	AQ707814	AQ707814 HS_5561_B
29	122	7.2	220	161	BB587413	BB587413 BB587413
30	122	7.2	470	166	BE326846	BE326846 hr65g04.x
31	121.2	7.2	780	18	AI317794	AI317794 uj20h04.y
32	118.4	7.0	914	106	AL531543	AL531543 AL531543
33	117.4	6.9	909	106	AL557971	AL557971 AL557971
34	115.4	6.8	980	17	AI226536	AI226536 uj09f09.y
35	112.6	6.7	788	16	AI121047	AI121047 ue89a11.y
36	112.2	6.6	819	16	AI114957	AI114957 ui41a04.y
37	111.8	6.6	817	24	AI746861	AI746861 ui08d04.y
38	110.8	6.5	773	16	AI119053	AI119053 ue93g04.y
39	109.8	6.5	897	111	AW106317	AW106317 um26b04.y
40	108	6.4	918	23	AI663558	AI663558 ui98c09.y
41	107.4	6.3	752	146	BF235733	BF235733 602026881
42	106.8	6.3	842	16	AI118969	AI118969 ue93b06.y
43	106.4	6.3	265	137	BE576686	BE576686 dc35d11.y
44	105	6.2	910	21	AI530076	AI530076 ui88h06.y
45	101.4	6.0	794	24	AI746972	AI746972 ul09g09.y

ALIGNMENTS

RESULT	1
AL546558	
LOCUS	AL546558 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI030YH12 5
DEFINITION	prime, mRNA sequence.
ACCESSION	AL546558
VERSION	AL546558.1 GI:12879791
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope

Genoscope - Centre National de Sequencage	
BP 191 91006 EVRY cedex - France	
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.	
Location/Qualifiers	
FEATURES	1..975
source	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="CS0DI030YH12"
	/clone_lib="LTI_NFL006_PL2"
	/tissue_type="placenta"
	/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT	311 a 228 c 215 g 220 t 1 others
ORIGIN	
Query Match 43.5%; Score 736; DB 106; Length 975;	
Best Local Similarity 96.8%; Pred. No. 4.2e-189;	
Matches 778; Conservative 0; Mismatches 10; Indels 16; Gaps 2;	
QY	874 tctctagagggaacaaagaaagaaatacaattcatgaattcctccattcactgcaaggctttt 933
Db	
QY	87 TATGCAGAGGGACAAAAGGAAAGAAAGAAATACAAATTCATGAATTCAAAAATCAGCAAAG 146
Db	
QY	934 actaccctaatacaaatagatccagcactgaagataaaacccaaaaagtgaatactgca 993
Db	
QY	147 ACTACCTTAATCAAAATAGATCCAGCACTGAAGATAAAACCAAAAAAGTGAATACTGCA 206
Db	
QY	994 gaccaatgtgctaataatagatgtactaggaataaaaggacttccattcactgcaaggctttt 1053
Db	
QY	207 GACCAATGTGCTAATAGATGTACTAGGAATAAAGGACTTCCATTCACTTGAAGGCTTTT 266
Db	
QY	1054 gtttttgataaaagcaagaaacaaatgcctctgtgtcccttcaatagcaagtgtgga 1113
Db	
QY	267 GTTTTGTATAAAGCAAGAAACAATGCCCTCTGGTTCCCTTCAATAGCATGTCAAGTGA 326
Db	
QY	1114 gtgaaaaagaatttgcccatgaatttgacctctatgaaacaaagactacattagaaac 1173
Db	
QY	327 GTGAAAAAGAATTTGGCCATGAATTGGACCTCTATGAAAAACAAGACTACATTAGAAAC 386
Db	
QY	1174 tgcataattggttaaaggacgcagctacaagggaacagtatctatcactaagagtggcatc 1233
Db	
QY	387 TGCATCATTTGGTAAAGGACCGCAGCTACAGGGAACAGTATCTATCACAAGAGTGGCATC 446
Db	
QY	1234 aaatgtcagccctggagttccatgataccacacgaacac-----agctat 1278
Db	
QY	447 AAATGTACGCCCTGGAGTTCCATGATACCACACGAACACAGCTTTTTCCTTCGAGCTAT 506
Db	
QY	1279 cggggtaaagacctacaggaaaactactgtcgaaatcctcgagggggaagggggaccc 1338
Db	
QY	507 CGGGTAAAGACCTACAGGAAACTACTGTGCGAAATCCTCGAGGGGAAGAGGGGACCC 566
Db	
QY	1339 tgggtttcacaaagcaatccagaggtacgctacgaagtctgtgacattcctcagtttca 1398
Db	
QY	567 TGGTGTTCACAAGCAATCCAGAGGTACGCTACGAAGTCTGTGACATTCCTCAGTGTTC 626
Db	
QY	1399 gaagtgaatgcatgacctgcaatggggagaggttatcgagggtctcatgatacacagaa 1458
Db	
QY	627 GAAGTTGAATGCATGACCTGCAATGGGAGAGTTATCGAGGTCTCATGGATCATACAGAA 686
Db	
QY	1459 tcaggcaagatttgtcagcgctgggatacatcacagaccacaccgacacaaattcttgcct 1518
Db	
QY	687 TCAGGCAAGATTGTTCAGCGCTGGGATCATCAGACACCACACCCGGCACAAATCTTTCCT 746
Db	
QY	1519 gaaagatattcccgacaagggttttgatgataattattgccgcaatcccgatggccagcg 1578
Db	

```

http://fulllength.invitrogen.com"
BASE COUNT      274 a   212 c   110 g   266 t   10 others
ORIGIN

Query Match      35.1%; Score 594.4; DB 106; Length 872;
Best Local Similarity 89.8%; Pred. NO. 1.3e-150;
Matches 666; Conservative 9; Mismatches 50; Indels 17; Gaps 3;

QY 874 tctctagaggacaaaagaagaagaataatacaattcatgaattcaaaaaaatcagcaaaag 933
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 132 TATTCATATTGACAAATTAATAATAATAATACAAATTCATTAATTCAAAAATCATCAAT 191

QY 934 actaccctaatacaaatagatccagcactgaagataaaaaacaaaagaagtgaatactgca 993
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 192 ACTACCCTAATCAAAATATATCCATCACTTAATATAAAACCAAAAAATTTAATACTGCA 251

QY 994 gaccaatgtcctaataagatgtactaggaataaaggacttccattcacttgcaaggctttt 1053
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 252 TACCAATTTTCTAATATATTTACTATTATAAATTAATTCATTCCATTCACTTCAAKKCTTT 311

QY 1054 gttttgataaagcaagaaaacaatgcctctgtgtcccttcaatagcatgtcaagtggga 1113
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 312 TTTTATTATAAATCAAGAAAACAATTCCTCTGTGTCCCTTCAATAGCAATTTCAAGTTGA 371

QY 1114 gtgaaaaaagaatttggccatgaatttgacctctatgaaaaacaaagactacattagaaac 1173
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 372 GTTAAAAAKAATTTGCCATTAAATTTACCTTATTAAACAAGACTACATTAKAAAC 431

QY 1174 tgcattcattggttaaaggacgcagctacaagggaacagtagtatctactaataagtgccatc 1233
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 432 TTCATCATATKGTAAAGACGCGAGCTACAAAGGGAACAGTAGTATCTACTAAGAGTGGCATC 491

QY 1234 aaatgtcagccctggagttccatgataccacagaacac-----agctat 1278
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 492 AAATGTGAGCCCTGGAGTTCATGATACCACACTAAACACATCTTTTTTTCCTTCTATCTAT 551

QY 1279 cggggtaaagacctacagaaaactactgtcgaatcctcagagggaagaaggggaccc 1338
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 552 CTKGGTAAAGACCTACAGGAAAACACTACTGTGGAATCCTCTAYGTGAATAMGKTGA-CC 610

QY 1339 tgggttttcacaagcaatccagaggttacgcgtacgaagtctgt-gacattcctcagtggtc 1397
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 611 TGGTGTTTCACAAGCAATCCAGAGGTACGCTACGAAGTCTGTGGACATTCCTCAGTGTTTC 670

QY 1398 agaagttgaaatgcatagcctgcaaatggggagagttatcgaggtctcatggatcatacaga 1457
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 671 AGAAGTTGAATGCATGACCTGCAATGGGGAGAGTTATCGAGGTCTCATGGATCATACAGA 730

QY 1458 atcaggcaagatttgcagcgctgggatcatcacacacccgcacaaattcttgcc 1517
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 731 ATCAGGCAAGATTGTGACGCGCTGGGATCATCACACACCACACCGGCACAAATTCCTGCC 790

QY 1518 tgaagatatcccgacaaggccttggatgataattattgccgcaatcccgatggccagcc 1577
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 791 TGAAAGATATCCCGACAAGGCGCTTTGATGATAATTATTGCCGCAATCCCGATGCCAGCS 850

QY 1578 gaggccatggtgctatactctt 1599
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 851 GAGGCCATGGTGCTATACTCTT 872

RESULT 4
AL546883
LOCUS      AL546883      828 bp      mRNA      EST      16-FEB-2001
DEFINITION AL546883 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI026YD16 5
prime, mRNA sequence.
ACCESSION AL546883
VERSION   AL546883.1 GI:12880433
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 828)
AUTHORS      Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE        Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      Contact: Genoscope
Genoscope - Centre National de Sequenage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
    Location/Qualifiers
        1..828
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="CS0DI026YD16"
            /clone_lib="LTI_NFL006_PL2"
            /tissue_type="placenta"
            /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      290 a   183 c   157 g   196 t
ORIGIN

Query Match      31.1%; Score 526.2; DB 106; Length 828;
Best Local Similarity 98.7%; Pred. No. 4.6e-132;
Matches 528; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 874 tctctagaggacaaaaggaaaagaataatacaattcatgaattcaaaaaaatcagcaaaag 933
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 96  TATGCAGAGGGACAAAGGAAAGAAAGAAATACAAATTCATGAATTCAAAAAATCAGCAAAG 155

QY 934 actaccctaatacaaatagatccagcactgaagataaaaaacaaaagaagtgaatactgca 993
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 156 ACTACCCTAATCAAAATAGATCCAGCACTGAAGATAAAACCAAAAGTGAATACTGCA 215

QY 994 gaccaatgtcctaataagatgtactaggaataaaggacttccattcacttgcaaggctttt 1053
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 216 GACCAATGTCTAATAGATGTACTAGGAATAAAGGACTTCCATTCACTTGAAGGCTTTT 275

QY 1054 gttttgataaagcaagaaaacaatgcctctgtgtcccttcaatagcatgtcaagtggga 1113
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 276 GTTTTGTATAAAGCAAGAAAACAATGCCTCTGTGTTCCCTTCAATAGCATGTCAAGTGA 335

QY 1114 gtgaaaaaagaatttggccatgaatttgacctctatgaaaaacaaagactacattagaaac 1173
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 336 GTGAAAAAAGAATTTGGCCATGAATTTGACCTCTATGAAAAACAAGACTACATTAGAAAC 395

QY 1174 tgcattcattggttaaaggacgcagctacaaggggaacagtagtatctactaagagtgccatc 1233
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 396 TGCATCATTTGGTAAAGGACGCGACTACAAGGGAACAGTATCTTWTCACTAAGAGTGGCATC 455

QY 1234 aaatgtcagccctggagttccatgataccacacgaacacagctatcggggtaaagacctta 1293
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 456 AAATGTCAGCCCTGGAGTTCATGATACACACGAACACAGCTATCGGGGTAAAGACCTTA 515

QY 1294 caggaaaactactgtcgaaatcctcaggggaagaaggggacccctggtgtttcacaaagc 1353
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 516 CAGGAAAACTACTGTGGAATCCTCGASGGGAAGAGGGGACCCCTGGTGTTCACAAAGC 575

QY 1354 aatccagaggtacgctacgaagtctgtgacattcctcagtggttcagaagttgaat 1408
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 576 AATCCAGAGGTACGCTACGAAGTCTGTGACATTCCCTCAGTGTTCAGAAGTAAAT 630

RESULT 5
BF383652
```

Db 61 GCC-GCTACAAGCGGTGGTGCCCGGCCCTTGGCAGGAGGATGTGSCAGATGCTGAAGAGT 119

QY 167 gtgctggtcgctgtgggccccttaattggaactgcccggcccttccactacaacagtgcagcc 226

Db 120 GTGCTGTGCTGTGGGCCCCCTTAATGGACTGCCGGCGGTCCACTACAATGTGACGACC 179

QY 227 atggttgccaaactgctgccatggactcaaacactgcgccccacagaggctgcgcttctg 286

Db 180 ATGTTGCCAACTGCTGCCATGGACTCAACACTCACCCACACAGAGGCTGCGGCATTCTG 239

QY 287 ggcgtgtgacctcttccagaagaagactacgtacggacctgcatcatgaacaatgggg 346

Db 240 GCGCTGTGACCTCTTCCAGGAGAAAG----- 266

QY 347 ttgggtaccggggcaccatggccacgacctgggtggcctgcccgtccaggcttgagcc 406

Db 267 ----- 266

QY 407 acaagttcccgaatgatcacaaagtatacacgcccactctccgggaatggcctggaagaaact 466

Db 267 -----GTACATGCCACCGCTCCGGATGGCCTGGAAGAGAACT 304

QY 467 tctgcgttaaccctgatggcgaccgccggaggtcccttggtgctacacaacagaccctgtg 526

Db 305 TCTGCGGTAAACCCTGATGGCGACCCCGGAGGTCTTGGTGCCACACAAACAGACCCTGCCG 364

QY 527 tgcgttccagagctcgggcatcaaatcctgcgggagggccgcgtgtgtctgtgcaatg 586

Db 365 TCGCTTCCAGAGCTGCGGCATCAATCTCTGCC-----GGCCGCGTGTGTGCTGCAATG 420

QY 587 gcgaggaataccgcggcggttagacgcgcaggtacgagtcagggcgagtgccagcgtggg 646

Db 421 GCGAGGAATACCGCGCGC-GTAGACCGCACCGAGTCAGGGCGCGAGTGCCAGCGCTGGG 479

QY 647 atcttcagcaccgcaccagcacccttcgagccgggcaagtctctcgaccaaggtctgg 706

Db 480 ATCTTCAGACCCCGCACCGACCCCTTCGAGCGGGCAAGTTCCTCGACCAAGGTCTGG 539

QY 707 acgacaactattgccggaatcctgacggctccgagcggccatggtgctacactacggatc 766

Db 540 ACGACAACTATTGCCGAATCCTGACGGCTCCGAGCGGCCATGGTCTACACTACGGATC 599

QY 767 cgcagatcgacgagagttctgtgacctcccccgctgcgggtccgaggtccaggtcccgcc 826

Db 600 CGCAGATCGACGAGAAATTCTGTGACCTCCCCCGTCCGGTAGCGCGGGGACGAGGCC 659

QY 827 tcgagggcggtggcggttctgtggtggcggtggtcctccggcggtggcg 872

Db 660 TGGGAGGTACCTGGGGAACCTTGGGGAGGGGGCGTGGCTTGGCCG 705

RESULT 7

BF234001

LOCUS BF234001 780 bp mRNA EST 14-NOV-2000

DEFINITION 602024489F1 NCI_CGAP_Li9 Mus musculus cDNA clone IMAGE:4159774 5', mRNA sequence.

ACCESSION BF234001

VERSION BF234001.1 GI:11145015

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 780)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM9438 row: a column: 23
High quality sequence stop: 768.

FEATURES

source Location/Qualifiers

1..780

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4159774"

/clone_lib="NCI_CGAP_Li9"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.9 kb. Constructed by Life Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 172 a 213 c 236 g 159 t

ORIGIN

Query Match 26.9%; Score 455.8; DB 146; Length 780;

Best Local Similarity 77.8%; Pred. No. 6.3e-113;

Matches 575; Conservative 0; Mismatches 162; Indels 2; Gaps 2;

QY 1 atgggggtgctcccaactcctgctcttctgactcaatgcttaggggtccctgggcagcgc 60

Db 44 ATGGGGTGGCTCCCACCTCTGCTGCTTCTGGTACAGTGTTCAGGGCTCTTGGGCAGCGC 103

QY 61 tcgccattgaatgacttccaaagtgcctccggggcacagagctacagcacctgctacatgcg 120

Db 104 TCACCACTGAATGACTTCCAGCTGTTCCGGGGCACAGAGTTAAGGAACCTGTTACACACA 163

QY 121 gtggtgccgggacctgacgaggaggtgtggtgagatgctgaagagtgtgctgctgctgt 180

Db 164 GCGGTGCCGGGCCATGGCAGGAGGATGTGGCAGATGCTGAGGAGTGTGCTAGGCGCTGT 223

QY 181 gggcccttaatggactgcggggcccttccactacaacgtgagcagccatggttgcgaactg 240

Db 224 GGGCCCCCTTCTGGACTGTGGGGCCTTCCACTACAACATGAGCAGCCATGGTTGCCAGCTG 283

QY 241 ctgccatggactcaacactgcggcccccacagaggtgcggttcttggcgctgtgacctc 300

Db 284 CTGCGGTGACCCAGCACCTCGCTGCACACACAGCTATACCCTCGAGTCTGTGCCATCTC 343

QY 301 ttccagaagaaagactacgtacgacctgcacatcatgaacaatgggttgggtaccggggc 360

Db 344 TTCCAGAAGAAAGATTATGTGGGACCTGCATTTATGGACAATGGGTGAGCTACCGGGGC 403

QY 361 accatggccacgacctggtggcctgcctgcctgcaggtgagccacaagtccccgaat 420

Db 404 ACTGTGGCCAGGACAGCTGTTGGCCTGCCCTGCCAAGCCTGGAGTCGCAGGTTCCCAAT 463

QY 421 gatcacaagtacacgcccactctccggaatggcctggaagaaacttctgccgtaaccct 480

Db 464 GACCACAAGTATACGCCCCACGCCAAAGAATGGCCTGGAAGAGAACTTCTGTAGGAACCCCT 523

QY 481 gatggcgaccccgagggtccttgggtgctacacaacagacctgctgctgcttccagagc 540

Db 524 GATGGGATCCCAGAGGTCCTGGTGTCTACACAACAACCCGAGTGTGCG-TTCCAGAGC 582

QY 541 tgcggcatcaaatcctgcgggagggcgctgtgtctgtgtggtgcaatggcaggaataccgc 600

Db 583 TGTGGCATCAAAACCTGCAGGAGGCTG-TTGTGTTCTGTGCAACGGTGAGGATTACCGT 641

QY 601 ggcgcggttagaccgcaggtcagggcgagtgccagcgcgtgggtatcttccagcaccgc 660

Db 642 GCGAGGTAGACGTTACAGAGTCAGGGCGGGAGTGTCAACCGTGGGACCCCTGCAGCACCC 701

QY 661 caccagcacccttcgagccgggcaagttcctcgaccaaggtctcggacgacaactattgc 720

Db 702 CAATCGCACCCCTTTCCAGCCTGAAAGTCTTTAGAAAAAGATCTGAAAGAACAACTATTGT 761


```

Query Match      25.3%; Score 428.4; DB 148; Length 652;
Best Local Similarity 81.3%; Pred. No. 1.7e-105;
Matches 509; Conservative 0; Mismatches 116; Indels 1; Gaps 1;

QY 1 atgggggtggtcccaactcctgctgcttctgactcaatgcttaggggtccctgggcagcgc 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 28 ATGGGGTGGCTCGCACTTCTGCTGCTTCTGGTACAGTGTTCAGAGGCTCTTGGCAGCGC 87

QY 61 tcgccattgaatgacttccaagtgtctccggggcacagagctacagcacctgtacatgcg 120
    || ||| ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 88 TCACCACTGAATGACTTCCAGCTGTTCCGGGGCACAGAGTTAAGGAACCTGTTACACACA 147

QY 121 gtggtgcccggcccttgccaggaggtgtggcagatgctgaagagtgtgctgctgctgt 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 148 GCGGTGCGGGGCCCATGGCAGGAGGATGTGGCAGATGCTGAGGAGTGTGCTAGGCGCTGT 207

QY 181 gggcccttaatggactgccgggccccttccactacaacgtgagcagccatggttgccaactg 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 208 GGGCCCCCTCTGGACTGTTCGGCCCTTCCACTACAACATGAGCAGCCATGGTTGCCAGCTG 267

QY 241 ctgccatggactcaacactgcgccccacacagaggctgcggcgttcttggtggcgctgtgacctc 300
    ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 268 CTGCCGTGGACCCAGCACTCGCTGCACACACAGCTATACCACTCGAGTCTGTGCCATCTC 327

QY 301 ttccagaagaaagactacgtacgcggacctgcacatcatgaacaaatgggttgggtaccggggc 360
    ||||| ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 328 TTCCAGAGAAGAAATATTGTGCGGACCCTGCATTTATGGACAATGGGGTTCAGCTACCGGGC 387

QY 361 accatggccacgaccgtgggtggcctgcctgccaggcttgagcagccaaagtccccaagt 420
    || ||||| ||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 388 ACTGTGGCAGGACAGCTGGTGGCCTGCCTGCCAAGCCTGGAGTCGAGGTTCCCAAT 447

QY 421 gatcacaagtacacgcccactctccggaatggcctggaagagaacttctgccgtaaccct 480
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 448 GACCACAAGTATACGCCACGCCAAAGAAATGGCCTGGAAGAGAACCTCTGTAGGAACCT 507

QY 481 gatggcagccccggagggttccttggtgctacacaaagaccctgctgtgcttccagagc 540
    ||||| || || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 508 GATGGGATCCCGAGAGGTCCCTGGTGTCTACACAAACCCGACGTGTCGTTTCCAGAGC 567

QY 541 tgcggcatcaaatcctgcgggagggcgctgtgtgtgctgtgcaatggcgagggaataaccgc 600
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 568 TGTGGCATCAAAACCTGCAGGGAGGCTG-TTGTGTTCTGTGCAACGGTGAGGATTACCGT 626

QY 601 ggcgcggttagaccgcacggagtcagg 626
    ||||| ||||| || ||||| |||||
Db 627 GCGAGGTAGACGTTACAGAGTCAAG 652

RESULT 10
AL546856
LOCUS      AL546856      844 bp      mRNA      EST      16-FEB-2001
DEFINITION      AL546856 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI026YL03 5
prime, mRNA sequence.
ACCESSION      AL546856
VERSION        AL546856.1 GI:12880379
KEYWORDS
SOURCE
ORGANISM      human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 844)
AUTHORS        Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE          Full-length cDNA libraries and normalization
JOURNAL        Unpublished (2001)
COMMENT        Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
source
1. .844
```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI026YL03"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT      301 a      152 c      57 g      304 t      30 others
ORIGIN

Query Match      24.5%; Score 414.2; DB 106; Length 844;
Best Local Similarity 83.6%; Pred. No. 1.3e-101;
Matches 460; Conservative 20; Mismatches 55; Indels 15; Gaps 1;

QY 874 tctctagagggacaaaaggaaagaagaatacaattcatgaattcaaaaaatcagcaaaag 933
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 14 TATGCAGAGGGACAAAGAAAGAAAGAAATACAATTTCATGAATTCAAAAAATCAGCAAAG 73

QY 934 actaccctaatcaaaatagatccagcactgaagataaaaaacccaaaagtgaatactgca 993
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 74 ACTACCTTAATCAAAATAGATCCAGCACTGAAGATAAAAAACCAAAAAAGTGAATCTGCA 133

QY 994 gaccaatgtgctaataatagatgtactaggaaataaaaggaccttccattcacttgcaaggctttt 1053
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 134 GACCAATGTGCTAATAGATGTACTAGGAATAAAGGACTTCCATTCTACTTGCAAGGCTTTT 193

QY 1054 gtttttgataaaagcaagaaacaatgcctctgtgttcccttcaatagcatgtcaagtga 1113
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 194 GTTTTGTATAAAGCAAGAAACAATGCCTCTGTGTTCCCTTCCAATAGCATGTCAAGTGGA 253

QY 1114 gtgaaaaaagaatttggccatgaatttgacctctatgaaacaaagactacattagaaac 1173
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 254 GTKAAAAAAGAATTTKKCCATGAATTTKACCTCTATKAAAACAAGACTACATTAGAAAC 313

QY 1174 tgcattgtgtaaaggacgcagctacaaagggaacagtatctatcactaagagtggcatc 1233
    ||||| ||||| ||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 314 TGCATCATTKGTAAAGGACKSAKSTACAATTTAASATTATCTATCATAATATTTTCATC 373

QY 1234 aaatgtcagccctggagttccatgatgataccacacgaacac-----agctat 1278
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 374 AAATTTTCATCCCTTTATTTTCCATTATACCACACTAACACATCTTTTTCSTTCTATCTAT 433

QY 1279 cggggtaaagacctacaggaataactactgtcgaatacctcgagggggaagaggggaccc 1338
    | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 434 CTTKTTAAATACCTACATTTAAAACTACTTTCTAAATCCTCTCTABTTTAATAMTTTTTACCC 493

QY 1339 tgggtgttcacaagcaatccagaggtacgctacggaagtctgtgacattcctcagtgttca 1398
    | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 494 TTTTTTTTCACAATCAATCCATATTACTCTACTAATNCTTTTASATTSSTSATTTTSTA 553

QY 1399 gaagttgaat 1408
    || |||||
Db 554 TAATTTAAAT 563

RESULT 11
BE723534
LOCUS      BE723534      543 bp      mRNA      EST      14-SEP-2000
DEFINITION      193407 MARC 4BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION      BE723534
VERSION        BE723534.1 GI:10124839
KEYWORDS
SOURCE
ORGANISM      Bos taurus
```


Db 497 AGGGGAAGAGGGGGACCCCTGGTGTTCACAAAGCAATCCAGAGGTACGCTACGAAGTCTG 438

QY 1380 tgacattcctcagtggttcagaaagttgaatgcatgacctgcaatgggagagttatcgagg 1439
|||||

Db 437 TGACATTCCCTCAGTGTTCAGAAAGTNGAATGCATGACCTGCAATGGGGAGAGTTATCGAGG 378

QY 1440 tctcatggtatcatacagaatcaggccaagatttgtcagcgctgggatcatcagacaccaca 1499
|||||

Db 377 TCTCAT-GATCATACAGAATCAGGCAAGA-TTGTGACGGCT-GGATCATCAGACACCACA 321

QY 1500 ccggcacaaaatttctgcctgaaagatatcccgacaagggttttgatgataattattgcgg 1559
|||||

Db 320 CCGGCACAAATCTTGCSTGAAAGATATCCCGACAAGGGCTTTTGATGATAATTATTGCCG 261

QY 1560 caatcccgatggccagccgagggccatggtgctatactcttgaccctcacaccccgctggga 1619
|||||

Db 260 CAATCCCGATGGCCAGCCGAGGCCATGGTGCTATACTCTTGACCCCTCACACCCGCTGGGA 201

QY 1620 gtactgtgcaattaaaacatgcg 1642
|||||

Db 200 GTACTGTGCAATTAANACATGCG 178

RESULT 13

AA037738

LOCUS AA037738 506 bp mRNA EST 01-FEB-1997

DEFINITION zf03b09.r1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone IMAGE:375833 5' similar to gb:M73239 HEPATOCYTE GROWTH FACTOR PRECURSOR (HUMAN);, mRNA sequence.

ACCESSION AA037738

VERSION AA037738.1 GI:1512912

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 506)

AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M. Generation and analysis of 280,000 human expressed sequence tags Genome Res. 6 (9), 807-828 (1996)

TITLE 97044478

JOURNAL Contact: Wilson RK

MEDLINE Washington University School of Medicine

COMMENT 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

Insert Length: 1135 Std Error: 0.00

Seq primer: -28M13 rev2 from Amersham

High quality sequence stop: 496.

Location/Qualifiers

1. .506

/organism="Homo sapiens"

/db_xref="GDB:1284089"

/db_xref="taxon:9606"

/clone="IMAGE:375833"

/clone_lib="Soares_fetal_heart_NbHH19W"

/sex="unknown"

/dev_stage="19 weeks"

/lab_host="DH10B (ampicillin resistant)"

/note="Organ: heart; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCGCATCTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into

source

the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbHL19W."

BASE COUNT 172 a 124 c 98 g 112 t

ORIGIN

Query Match 20.7%; Score 350.2; DB 1; Length 506;

Best Local Similarity 98.9%; Pred. No. 2.9e-84;

Matches 363; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 874 tctctagagggacaaaaggaagaagaatacataattcatgaattcaaaaaatcagcaaaag 933
| | | | |

Db 140 TATGCAGAGGGACAAAAGGAAAAGAAATAACAATTTCATGAATTCAAAAATCAGCAAAG 199

QY 934 actaccctaatacaaatagatccagcactgaagataaaaaaaccaaaagtgaatactgca 993
| | | | |

Db 200 ACTACCCTAATCAAAATAGATCCAGACTGAAGATAAAAAACCAAAAAAGTGAATACTGCA 259

QY 994 gaccaatgtgctaataatagatgtactagggaataaaaggacttccattcacttgcaaggctttt 1053
| | | | |

Db 260 GACCAATGTGCTAATAGATGTACTAGGAATAAAGGACTTCCATTCCATTGCAAGGCTTTT 319

QY 1054 gtttttgataaaagcaagaaaaaataatgcctctgtgttcccttcaatagcatgtcaagtggga 1113
| | | | |

Db 320 GTTTTGTATAAAGCAAGAAAACAATGCCTCTGTGTTCCCTTCAATAGCATGTCAAGTGGGA 379

QY 1114 gtgaaaaaagaatttggccatgaatttgacctctatgaaacaaaagactacattagaaac 1173
| | | | |

Db 380 GTGAAAAAAGAATTTGGCCATGAATTTGACCTCTATGAAAAACAAGACTACATTAGAAAC 439

QY 1174 t-gcatcattggttaaaggacgcagctacaaggggaacagtatctatcactaagagtggtcat 1232
| | | | |

Db 440 TGGCATCATTTGGTAAAGGACGCGCAGCTACAAAGGGAACAGTATCTATCACTAAGAGTGGCAT 499

QY 1233 caaatgt 1239

Db 500 CAAATGT 506

RESULT 14

BE755132

LOCUS BE755132 518 bp mRNA EST 15-SEP-2000

DEFINITION 208867 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION BE755132

VERSION BE755132.1 GI:10169124

KEYWORDS EST.

SOURCE COW.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae; Bovidae; Bovinae; Bos. 1 (bases 1 to 518)

REFERENCE Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Laegreid ,W.W. and Keele,J.W.

AUTHORS Design and use of four pooled tissue normalized cDNA libraries for EST discovery in cattle

TITLE Unpublished (2000)

JOURNAL Contact: Smith TPL

COMMENT USDA, ARS, US Meat Animal Research Center

PO Box 166, Clay Center, NE 68933-0166, USA

Tel: 402 762 4366

Fax: 402 762 4390

Email: smith@email.marc.usda.gov

Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.

PCR Primers

FORWARD: AGGAAACAGCTATGACCAT

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2001, 16:21:43 ; Search time 4515.12 Seconds
(without alignments)
5909.448 Million cell updates/sec

Title: US-09-600-991-1
Perfect score: 1725
Sequence: 1 atgtggtgaccaaactoct.....accaccaccaccaccactag 1725

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vil:*
59: gb_vi2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rol:*
95: gb_ro2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1725	100.0	1725	9	AX019527	AX019527 Sequence
2	996.6	57.8	1692	9	AX019529	AX019529 Sequence
3	857.8	49.7	2190	9	AX019515	AX019515 Sequence
4	857.8	49.7	2190	9	AX019517	AX019517 Sequence
5	857.8	49.7	2199	10	E05008	E05008 HLC2 CDNA S
6	857.8	49.7	2199	10	E16560	E16560 Human HLC3
7	857.8	49.7	2289	10	E12582	E12582 CDNA encodi
8	857.8	49.7	2704	97	HUMSCFA2	M73240 Human (clon

QY	1141	agtgagtgaaaaagaatttggccatgaatttgacctctatgaaacaaagactacatt	1200
Db	1141	AGTGGAGTGAAAAAGAAATTGGCCATGAATTGACCTCTATGAAAAACAAGACTACATT	1200
QY	1201	agaaactgcatcatgtggtaaaggacgcagctacaagggaacagtagtatctatactaagagt	1260
Db	1201	AGAAACTGCATCATTTGTTAAAGGACGCAGCTACAAGGAACAGTATCTATCACTAAGAGT	1260
QY	1261	ggcatcaaatgtcagccctggagttccatgataccacacgaacacagctatcggggtaaa	1320
Db	1261	GGCATCAAATGTCAGCCCTGGAGTTCCATGATACCACACGAACACAGCTATCGGGGTAAA	1320
QY	1321	gacctacagggaaaactactgtcgaaaatcctcgagggggaagaaggggacccctggtgttc	1380
Db	1321	GACCTACAGGAAAACACTACTGTGCGAAATCCTCGAGGGGAAGAAGGGGACCCCTGGTGTTC	1380
QY	1381	acaagcaatccagaggtacgctacgaagtctgtgacattcctcagtgttcagaagttgaa	1440
Db	1381	ACAAGCAATCCAGAGGTACGCTACGAAGTCTGTGACATTCTCAGTGTTCAGAAGTTGAA	1440
QY	1441	tgcatgacctgcaatggggagaggttatcgaggtctcatggtatcatcacagaatcagggaag	1500
Db	1441	TGCATGACCTGCAATGGGGAGAGTTATCGAGGTCTCATGGATCATACAGAATCAGGCAAG	1500
QY	1501	attgtcagcgctgggatcatcagacacacacccggcacaaaattcttgctgaaagatat	1560
Db	1501	ATTGTGTCAGCGCTGGGATCATCAGACACCAACCGGCACAAATTCTTGCCCTGAAAGATAT	1560
QY	1561	cccgaacaaggctttgatgataaattattgcccgaatcccgatggccagccgagggccatgg	1620
Db	1561	CCCGACAAGGGCTTTGATGATAATTATTGCCGGAATCCCGATGGCCAGCCGAGCCCATGG	1620
QY	1621	tgctatactcttgacctcacaccccgctggggagtactgtgcaattaaacacatgcgctgac	1680
Db	1621	TGCTATACTCTTGACCTCACACCCGCTGGGAGTACTGTGCAATTAAACATGCGCTGAC	1680
QY	1681	aaagctgacgacgacgacaaaacaccaccaccaccaccaccactag	1725
Db	1681	AAAGCTGACGACGACGACAAACACCACCACCACCACCACCACCACCTAG	1725
RESULT	2		
AX019529			
LOCUS	AX019529	1692 bp	DNA
DEFINITION	Sequence 3 from Patent WO9938967.		
ACCESSION	AX019529		
VERSION	AX019529.1	GI:10043446	
KEYWORDS			
SOURCE	synthetic construct.		
ORGANISM	synthetic construct		
REFERENCE	artificial sequence.		
AUTHORS	1 (bases 1 to 1692)		
TITLE	Medico,E., Collesi,C., Comoglio,P., Michieli,P. and Caselli,G.		
JOURNAL	Recombinant proteins derived from hgf and msp		
	Patent: WO 9938967-A 3 05-AUG-1999;		
	DOMPE SPA (IT); MEDICO ENZO (IT); COLLESI CHIARA (IT); COMOGLIO		
	PAOLO (IT); MICHIELI PAOLO (IT); CASELLI GIANFRANCO (IT)		
FEATURES	Location/Qualifiers		
source	1..1692		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
	/notes="RECOMBINANT"		
BASE COUNT	442 a	457 c	459 g 334 t
ORIGIN			
Query Match	57.8%; Score 996.6; DB 9; Length 1692;		
Best Local Similarity	83.2%; Pred. NO. 1.6e-255;		
Matches 1166;	Conservative	0; Mismatches 214;	Indels 21; Gaps 2;
QY	346	tttgacctctatgaaacaaagactacattagaaactgcatcattggttaaagcagcagc	405
Db	1		

Db	292	TGTGACCTCTTCCAGAAGAAAGACTACGTACGGACCTGCATCATGAACAATGGGTTGGG	351
QY	406	tacaagggaaacagtatctatcataaagatgggcatcaaatgtcagccctggagttccatg	465
Db	352	TACCGGGGCACCATGGCCACGACCGTGGTGGCTGCCCTGCCAGGCTTGGAGCCACAAG	411
QY	466	ataccacacgaacacacagcta-----tcggggtaaaagacctacagggaaaactactgt	516
Db	412	TTCCCGAATGATCAACAAGTACACGCCCACTCTCCGGAATGGCCTGGAAGAACTTCTGC	471
QY	517	cgaaatcctcgaggggaagaaggggacccctggtgtttcacagaacaatccagaggtacgc	576
Db	472	CGTAACCCCTGATGGCGACCCCGAGGTCTTGGTGTACACAACAGACCCCTGCTGTGCGC	531
QY	577	tacgaagtctgtgacattcctcagtggttcagaagttgaatgcatgacctgcaatggggag	636
Db	532	TTCAGAGCTGCGGCATCAAATCCTCCGGGAGGCGCGTGTCTCTGGTGCAATGGCGAG	591
QY	637	agttatcgaggtctcatggatcatacagaatcaggcaagatttgcagcgctgggatcat	696
Db	592	GAATACCGCGCGGTAGCCCGACGGAGTCAAGGCGGAGTGCACGCGCTGGGATCTT	651
QY	697	cagacaccacacccggcacaaaattctgcctgaaagatatccgcacaagggctttgatgat	756
Db	652	CAGACCCGACCCAGCACCCCTTCGAGCCGGGCAAGTTCCTCGACCAAGGTCTGGACGAC	711
QY	757	aattattgccgcaatcccgatggccagcccgaggccatggtgctatactcttgacccctcac	816
Db	712	AACTATTGCCGGAATCCTGACGGCTCCGAGCGGCCATGGTGCTACACTACGGATCCGCAG	771
QY	817	accgctgggagtagtactgtgcaattaaaaacatgcgctgacaaagc-----ttcg	864
Db	772	ATCGAGCGAGAGTTCTGTGACCTCCCGCGTGGGTCCGAGGCACAGCCCGCCTCGAG	831
QY	865	ggcggtggcggttctggtggcggtggctccggcggtggtggttctctagagggacaaagg	924
Db	832	GCGGTGGCGGTTCTGGTGGCGGTGCTCCGCGGTGGCGGTTCTCTAGAGGGACAAAGG	891
QY	925	aaaagaagaataacaaattcatgaattcaaaaaaatcagcaaaagactaccctaataata	984
Db	892	AAAAGAAGAAATACAAATTCATGAATTCAAAAAATCAGCAAAAGACTACCTTAATCAAATA	951
QY	985	gattcagcactgaagataaaaaaccataatgcatgcaagtggagtgagtaaaaaagaattggc	1044
Db	952	GATCCAGCACTGAAGATAAAAAACCACAAAAGTGAATACTGCAGACCAATGTGTAATAGA	1011
QY	1045	tgtactaggaataaaggacttccattccacttgcaaggctttgtttttgataaaagcaaga	1104
Db	1012	TGTA TAGGAATAAAGGACTTCCATTCACTTGCAAGGCTTTTGTGTTTGTATAAAGCAAGA	1071
QY	1105	aaacaatgcctctggttcccttcaatagcatgtcaagtggagtgagtaaaaaagaattggc	1164
Db	1072	AAACAATGCCTCTGGTTCCCTTCAATAGCATGTCAAGTGGAGTGAAAAAAGAATTGGC	1131
QY	1165	catgaatttgacctctatgaaaacaaagactacattagaaactgcatcattggtaaagga	1224
Db	1132	CATGAATTTGACCTCTATGAAAAACAAAGACTACATTAGAAACTGCATCATTTGGTAAAGGA	1191
QY	1225	cgcagctacaagggaacagtagtatctatacctaagagtggtgcatcaaatgtcagccctgga	1284
Db	1192	CGCAGCTACAAGGGGAACAGTATCTATCACTAAGAGTGGCATCAAAATGTCAGCCCTGGAGT	1251
QY	1285	tccatgataccacacgaacacagctatcggggttaaagacccctacagggaaaactactgtcga	1344
Db	1252	TCCATGATACCACACGAACACAGCTATCGGGGTAAAGACCTACAGGAAAACTACTGTCTGA	1311
QY	1345	aatcctcgaggggaagaagggggacccctggtgtttcacagaacaatccagaggtacgctac	1404
Db	1312	AATCTCGAGGGGAAGAAGGGGACCCCTGGTGTTCACAAGCAATCCAGAGGTACGTAC	1371
QY	1405	gaagtctgtgacattcctcagtggttcagaagttgaaatgcatgacctgcaatggggagagt	1464
Db	1372	GAAGTCTGTGACATTCCTCAGTGTTCAGAAGTTGAATGCATGCATGCACCTGCAATGGGAGAGT	1431


```
|||||
Db 61 CTGCTCCCATCGCCATCCCTATGCGAGGGGACAAAGGAAAGAAATACAATTCAAT 120
QY 121 gaattcaaaaaaatcagcaaaagactaccctaatacaaaaatagatccagcactgaagataaaa 180
Db 121 GAATTCAAAAAAATCAGCAAAGACTACCCCTAATCAAAAATAGATCCAGCACTGAAGATAAAA 180
QY 181 accaaaaagtgaatactgcagaccaatgtgctaataagtagatgtactaggaataaaggactt 240
Db 181 ACCAAAAAAGTGAATACTGCAGACCAATGTGCTAATAGATGTACTAGGAATAAAGGACTT 240
QY 241 ccattcacttgcaaggctttgttttgataaaagcaagaaaaacaatgcctctggttcccc 300
Db 241 CCATTCTACTTGCAAGGCTTTTGTTTTGATAAAGCAAGAAACAATGCCTCTGGTTCCCC 300
QY 301 ttcaatagcatgtcaagtggagtgaaaaaagaatttggccatgaatttgacctctatgaa 360
Db 301 TTCAATAGCATGTCAAGTGGAGTGAAAAAAGAATTTGGCCCATGAATTTGACCTCTATGAA 360
QY 361 aacaaagactacattagaaactgcatcatttgtaaaaggacgcagctacaagggaaacagta 420
Db 361 AACAAAGACTACATTAGAAACTGCATCATTTGTTAAAGGACGCAGCTACAAGGGAAACAGTA 420
QY 421 tctatactaagagtggcatcaaatgtcagccctgaggttccatgataccacacgaacac 480
Db 421 TCTATCACTAAGAGTGGCATCAATGTCAAGCCCTGAGCTGGAGTTCCATGATACCACACGAAC 480
QY 481 agctatcggggtaaagacctaagcaaaaactactgtcgaaatcctcgaggggaagaagg 540
Db 481 AGCTATCGGGGTAAAGACCTACAGSAAAACTACTGTGAAATCCTCGAGGGGAAGAGG 540
QY 541 ggaccttggtgtttcacaaagcaatccagaggttacgctacgaagtctgtgacattcctcag 600
Db 541 GGACCTTGGTGTTCACAAGCAATCCAGAGGTACGCTACGGAAGTCTGTGACATTCCTCAG 600
QY 601 tgttcagaagttgaatgcatgacctgcaatgggagagttatcgaggtctcatggatcat 660
Db 601 TGTTTCAGAAGTTGAATGCATGACCTGCAATGGGAGAGTTATCGAGGTCTCATGGATCAT 660
QY 661 acagaatcaggcaagatttgtcagcgctgggatcatcagacacacaccggcacaaattc 720
Db 661 ACAGAAATCAGGCAAGATTGTGACGCTGGGATCATCAGACACCACACCCGGCACAAATTC 720
QY 721 ttgcctgaaagatatcccgacaaaggcttttgatgataattattgcgcaatcccgatggc 780
Db 721 TTGCCTGAAAGATATCCCGACAAAGSGCTTTGATGATAATTATTGCGCAATCCCGATGGC 780
QY 781 cagccgagccatggtgctatactcttgacccctcacaccccgctgggagtagtctgcaatt 840
Db 781 CAGCCGAGCCCATGGTGCTATACTCTTGACCCCTCACACCCCGCTGGGAGTACTGTGCAATT 840
QY 841 aaaacatgcgctgacaaaagct 861
Db 841 AAAACATGCGCTGACAATACT 861

RESULT 5
E05008
LOCUS E05008 2199 bp RNA PAT 29-SEP-1997
DEFINITION HLC2 cDNA sequence coding for human hepatoparenchymal cell growth factor(HGF).
ACCESSION E05008
VERSION E05008.1 GI:2173203
KEYWORDS JP 1993111383-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2199)
AUTHORS Nakamura,T., Hagiya,M., Seki,T., Shimonishi,M., Shimizu,S.,
Inohara,I., Sakaguchi,M. and Asami,O.
TITLE RECOMBINANT HUMAN HEPATOCYTE GROWTH FACTOR AND ITS PRODUCTION
JOURNAL Patent: JP 1993111383-A 2 07-MAY-1993;
```

```
NAKAMURA TOSHIICHI
OS Homo sapiens (human)
PN JP 1993111383-A/2
PD 07-MAY-1993
PF 06-JUN-1991 JP 1991163485
PR 11-JUN-1990 JP 90P 152474
PI NAKAMURA TOSHIICHI, HAGIYA MICHIO, SEKI TATSUYA, PI
SHIMONISHI MANABU,
PI SHIMIZU SHIN, INOHARA IZUMI, SAKAGUCHI MARIKO, ASAMI OSAMU PC
C12N15/16,C07K13/00,C12N5/10,C12N15/85,C12P21/02//A61K37/02, PC
A61K37/24,
PC (C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC *source: tissue_type=Blood;
CC *source: cell_type=Leukocyte;
FH key Location/Qualifiers
FH 5'UTR 1..17
FT sig_peptide 18..110
FT /product='human HGF signal peptide' FT CDS
FT 111..2172 /product='human HGF'.
FEATURES
source Location/Qualifiers
1..2199
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 678 a 464 c 505 g 552 t
ORIGIN
Query Match 49.7%; Score 857.8; DB 10; Length 2199;
Best Local Similarity 99.8%; Pred. No. 2e-218;
Matches 859; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 atgtgggtgaccaaaactcctgccagccctgctgctgcagcatgtcctcctgcctcctc 60
Db 28 ATGTGGGTGACCAAACTCTGCCAGCCCTGCTGTGTCGAGCATGTCTCTCTGTCATCTCCTC 87
QY 61 ctgtcccccctgcgcataccctctatgcagagggacaaaagaaaagaaatacaattcat 120
Db 88 CTGCTCCCCATCGCCATCCCTATGCGAGGGACAAAGGAAAGAAAGAAATACAATTCAAT 147
QY 121 gaattcaaaaaatcagcaaaagactaccctaatacaaaatagatccagcactgaagataaaa 180
Db 148 GAATTCAAAAAATCAGAAAGACTACCCCTAATCAAAATAGATCCAGCACTGAAGATAAAA 207
QY 181 accaaaaaagtgaatactgcagaccaatgtgctaataagtagtactaggaataaaggactt 240
Db 208 ACCAAAAAAGTGAATACTGCAGACCAATGTGCTAATAGATGTACTAGGAATAAAGGACTT 267
QY 241 ccattcacttgcaaggcttttgttttgataaaagcaagaaaaacaatgcctctggttcccc 300
Db 268 CCATTCACTTGCAAGGCTTTTGTTTTGATAAAGCAAGAAACAATGCCTCTGTTCCCC 327
QY 301 ttcaatagcatgtcaagtggagtgaaaaaagaatttggccatgaatttgacctctatgaa 360
Db 328 TTCAATAGCATGTCAAGTGGAGTGAAAAAAGAATTTGGCCATGAATTTGACCTCTATGAA 387
QY 361 aacaaagactacattagaaaactgcattgtgtaaaaggcagcagctacaagggaaacagta 420
Db 388 AACAAAGACTACATTAGAAACTGCATCATTTGGTAAAGGACGCAGCTACAAGGGAACAGTA 447
QY 421 tctatcactaagagtggcatcaaatgtcagccctgaggttccatgatataccacacgaacac 480
Db 448 TCTATCACTAAGAGTGGCATCAAAATGTGAGCCCTGGAGTTCCATGATACCACACGAACAC 507
QY 481 agctatcggggtaaagacctacaggaanaactactgtcgaaaactcctcgaggggaagaagg 540
Db 508 AGCTATCGGGTAAAGACCTACAGGAAACTACTGTGAAATCCTCGAGGGGAAGAGGG 567
QY 541 ggaccttggtgtttcacaaagcaatccagaggttacgctacgaagtctgtgacattcctcag 600
|||||
```


QY	241	ccattcacttgcaaggcttttgtttttgataaagcaagaaaaacaatgcctctggttcccc	300
Db	241		
QY	301	ttcaatagcatgtcaagtggagtgaataaagaatttggccatgaatttgacctctatgaa	360
Db	301		
QY	361	aacaaagactacattagaaactgcatcattgttaaaggacgcagctacaagggaacagta	420
Db	361		
QY	421	tctatcactaagagtgccatcaaatgtcagccctggagttccatgataccacgaacac	480
Db	421		
QY	481	agctatcgggttaaagacctacaggaataactactgtcgaatacctcgcaggggaagagg	540
Db	481		
QY	541	ggacctggtgtttcacaaagcaatccagaggtacgctacgaagtctgtgacattcctcag	600
Db	541		
QY	601	tgttcagaagtgaatgcatgacctgcaatggggagagttatcgagggtctcatggatcat	660
Db	601		
QY	661	acagaatcaggcaagatttgtcagcgtgggatcatcagacaccacaccgggcacaaattc	720
Db	661		
QY	721	ttgcctgaaagatatcccgacaaaggcttttgatgataaattattgcccgaatcccgatggc	780
Db	721		
QY	781	cagccgaggccatggtgctatactcttgacctcacaccctcgcggagtagtctgtgcaatt	840
Db	781		
QY	841	aaaacatgcgctgacaaaagct	861
Db	841		
RESULT 10			
LOCUS	I55856	2288 bp	DNA
DEFINITION	Sequence 4 from patent US 5648233.		
ACCESSION	I55856		
VERSION	I55856.1	GI:2476650	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2288)		
AUTHORS	Yamaguchi,K., Shima,N., Murakami,A., Goto,M., Tsuda,E., Masunaga,H., Takahira,R., Oogaki,F., Ueda,M. and Higashio,K.		
TITLE	Modified tumor cytotoxic factor (TCF) and DNA encoding such		
JOURNAL	Patent: US 5648233-A 4 15-JUL-1997;		
FEATURES	Location/Qualifiers		
source	1..2288		
BASE COUNT	702 a	491 c.	524 g 571 t
ORIGIN	/organism="unknown"		
Query Match 49.0%; Score 845.8; DB 10; Length 2288;			
Best Local Similarity 99.7%; Pred. No. 3.3e-215;			
Matches 858; Conservative 0; Mismatches 2; Indels 1; Gaps 1;			
QY	1	atgtgggtgacaaactcctgcccagccctgctgctgcagcatgctcctcctgcattcctc	60

Db	78	ATGTGGGTGACCAAACTCTGTGCCAGCCCTGTCTGTGCAGCATGTCTCTCTGCATCTCCTC	137
QY	61	ctgtcccccatcgccatccccctatgcagaggggacaaaagaaaagaaatacaaatcat	120
Db	138	CTGCTCCCCCATCGCCATCCCCTATGCAGAGGGACAAAGGAAAGAAATACAATT	197
QY	121	gaattcaaaaaaatcagcaaaagactaccctaatcaaaaatagatccagcactgaagataaaa	180
Db	198	GAATTCAAAAAATCAGCAAAAGACTACCCTAATCAAAA-AGATCCAGCCTGAAGATAAAA	256
QY	181	acaaaaaagtgaatactgcagaccaaatgtgctaataatagatgtactaggaataaaggactt	240
Db	257	ACCAAAAAAGTGAATACTGCAGACCAATGTCTAATAGATGTACTAGGAATAAAGGACTT	316
QY	241	ccattcacttgcaaggcttttgtttttgataaaagcaaaaaacaatgcctctggttcccc	300
Db	317	CCATTCACTTGAAGGCTTTTGTTTTGATAAAAGCAAGAAACAATGCCTCTGGTTCCCC	376
QY	301	ttcaatagcatgtcaagtggagtgaataaagaatttggccatgaatttgacctctatgaa	360
Db	377	TTCAATAGCATGTCAAGTGGAGTGAAAAAGAAATTTGGCCATGAATTTGACCTCTATGAA	436
QY	361	aacaaagactacattagaaaactgcatcatttggtaaaggacgcagctacaagggaacagta	420
Db	437	AACAAAGACTACATTAGAAACTGCATCATTTGGTTAAAGGACGCAGCTACAAGGGAACAGTA	496
QY	421	tctatcactaagagtgcatcaaatgtcagccctgagctcagaaatcctcgcaggggaaggg	540
Db	497	TCTATCACTAAGAGTGGCATCAAAATGTGAGCCCTGGAGTTCCATGATACACACGAACAC	556
QY	481	agctatcggggtaaagacctacagaaaaactactgtcgaatacctcgcaggggaaggg	540
Db	557	AGCTATCGGGGTAAAGACCTACAGGAAAACTACTGTGAAAATCCTCGAGGGGAAAGGG	616
QY	541	ggacctggtgtttcacaaagcaatccagaggtacgctacgaagtctgtgacattcctcag	600
Db	617	GGACCTGGTGTTCACAAGCAATCCAGAGGTACGCTACGAAGTCTGTGACATTCCTCAG	676
QY	601	tgttcagaagttgaatgcatgacctgcaatggggagagttatcgagggtctcatggatcat	660
Db	677	TGTTCAGAAAGTTGAATGCATGACCTGCAATGGGGAGAGTTATCGAGGTCTCATGGATCAT	736
QY	661	acagaatcaggcaagatttgtcagcgtgggatcatcagacaccacaccggcacaaattc	720
Db	737	ACAGAAATCAGGCAAGATTTGTACGCGCTGGGATCATCAGACACCACCCGGCACAAATTC	796
QY	721	ttgcctgaaagatatcccgacaaaggcttttgatgataaattattgcccgaatcccgatggc	780
Db	797	TTGCCTGAAAAGATATCCCGACAAAGGGCTTTGATGATAATTTATTCGCCAATCCCGATGGC	856
QY	781	cagccgaggccatggtgctatactcttgacctcacaccctcgcggagtagtctgcaatt	840
Db	857	CAGCCGAGGCCATGGTGTATATCTTGACCCCTCACACCCCGCTGGGAGTACTGTGCAATT	916
QY	841	aaaacatgcgctgacaaaagct	861
Db	917	AAACATGCGCTGACAATACT	937
RESULT 11			
LOCUS	E06576	2184 bp	RNA
DEFINITION	cDNA encoding human growth factor for endothelial cell.		
ACCESSION	E06576		
VERSION	E06576.1	GI:2174763	
KEYWORDS	JP 1994009691-A/1.		
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 2184)		
AUTHORS	Shudo,T., Harada,K., Hirahara,I. and Adachi,M.		
TITLE	PROTEIN ORIGINATED FROM HUMAN		

Db 61 CTGCTCCCCATCGCCATCCCCTATGCAGAGGGACAAAGGAAAGAAATACAAATTCAT 120
QY 121 gaattcaaaaaatcagcaaaagactaccctaatacaaaaatagatccagcactgaagataaaa 180
Db 121 GAATTCAAAAATCAGCAAAAGACTACCCCTAATCAAAATAGATCCAGCAGCTGAAGATAAAA 180
QY 181 accaaaaaagtgaatactgcagaccacaaatgtgcttaataatagatgtactaggaataaaagactt 240
Db 181 ACCAAAAAAGTGAATACTGCAGACCAATGTGCTAATAATAGATGTACTAGGAATAAAGGACTT 240
QY 241 ccattcacttgcaaggcttttgttttgataaaagcaagaaaaaacaatgcctctggttcccc 300
Db 241 CCATTCACTTGCAAGGCTTTTGTTTTGATAAAGCAAGAAAAACAATGCCCTCTGTTCCCC 300
QY 301 ttcaatagcatgtcaagtggagtgaaaaaagaatttggcccatgaatttgacctctatgaa 360
Db 301 TTCAATAGCATGTCAAGTGGAGTGAAAAAAGAATTGGCCCATGAATTTGACCTCTATGAA 360
QY 361 acaaaagactacattagaaaactgcatacttggttaaggacgcagctacaaggggaacagta 420
Db 361 AACAAAGACTACATTAGAAAACCTGCATCATTTGGTAAAGGACGCAGCTACAAGGGAAACAGTA 420
QY 421 tctatcactaagtggtgcacaaatgtcagccctgggtttcacaaagcaatccagaggtacgctacgaacac 480
Db 421 TCTATCACTAAGAGTGGCATCAAATGTGAGCCCTGGAGTTCCATGATACCAACACGAAACAC 480
QY 481 -----agctatcgggggtaaaagacctacaggaataactactgtcgaaatcct 525
Db 481 AGCTTTTTCCTTCGAGCTATCGGGGTAAAGACCTACAGGAAAACCTACTGTGCGAAATCCT 540
QY 526 cgaggggaagaaggggaccctgggtttcacaaagcaatccagaggtacgctacgaagtc 585
Db 541 CGAGGGGAAGAAGGGGACCCTGGTGTTCACRAGCAATCCAGAGGTACGCTACGGAAGTC 600
QY 586 tgtgacattcctcagtttcagaagttgaatgcaatgcacctgcaatggggagagttatcga 645
Db 601 TGTGACATTCTCCTCAGTGTTCAGAAAGTTGAATGCATGACCTGCAATGGGGAGAGTTATCGA 660
QY 646 ggtctcatggatcatacagaatcaggcaagatttgtcagcgtgggatcatcagacacca 705
Db 661 GGTCTCATGGATCATACAGAATCAGGCAAGATTGTGTCAGCGCTGGGATCATCAGACACCA 720
QY 706 caccggcacaaattctgcctgaaagatatcccgacaaggccttggatgataaattatgc 765
Db 721 CACCGGCACAAATTCTGCCTGAAAGATATCCCGACAAGGGCTTTTGATGATAATTATTCG 780
QY 766 cgcaatcccgatggccagccgagccatggtgctatactcttgacccctcacacccgctgg 825
Db 781 CGCAATCCCGATGGCCAGCCGAGGCCATGGTGCTATACTCTTGACCCCTCACACCCGCTGG 840
QY 826 gagtactgtgcaattaaaaacatgcgctgacaaaagct 861
Db 841 GAGTACTGTGCAATTAAACATGCGCTGACAAATACT 876

RESULT 13
E09626
ID E09626 standard; RNA; HUM; 2184 BP.
XX
AC E09626;
XX
SV E09626.1
XX

07-OCT-1997 (Rel. 52, Created)
DT 02-SEP-2000 (Rel. 65, Last updated, Version 2)
XX
DE DNA encoding human HGF.
XX
KW JP 1995179356-A/1.
XX

OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN [1]
RP 1-2184
RA Nakamura T., Matsumoto K.;
RT "EPITHELIOCYTE GROWTH PROMOTER";
RL Patent number JP1995179356-A/1, 18-JUL-1995.
RL NAKAMURA TOSHIICHI.
XX
CC OS Homo sapiens (human)
CC PN JP 1995179356-A/1
CC PD 18-JUL-1995
CC PF 28-DEC-1990 JP 1990419158
CC PI NAKAMURA TOSHIICHI, MATSUMOTO KUNIO
CC PC A61K38/00,A61K38/00//C12N15/16,C12P21/02,(C12P21/02,C12R1:91);
CC CC strandedness: Double;
CC CC topology: Linear;
CC FH Key Location/Qualifiers
CC FH
CC FT source 1. .2184
CC FT /organism="Homo sapiens"
CC FT /tissue_type="liver"
CC FT
CC FT mat_peptide 163. .1482
CC FT /product="HGF alpha chain"
CC FT mat_peptide 1483. .2184
CC FT /product="HGF beta chain"
XX
FH Key Location/Qualifiers
FH
FH FT source 1. .2184
FH FT /db_xref="taxon:9606"
FH FT /organism="Homo sapiens"
XX
SQ Sequence 2184 BP; 673 A; 455 C; 500 G; 556 T; 0 other;

Query Match 48.3%; Score 832.8; DB 45; Length 2184;
Best Local Similarity 98.1%; Pred. No. 9.7e-212;
Matches 859; Conservative 0; Mismatches 2; Indels 15; Gaps 1;

QY 1 atgtgggtgaccaaactcctgccagccctgctgctgcagcatgctcctcctgcctcctc 60
Db 1 ATGTGGGTGACCAAACTCCTGCCAGCCCTGCTGCTGCAGCATGTCCTCCTGCATCTCCTC 60
QY 61 ctgctccccatgcctatcccctatgcagaggggacaaaaagaaagaataataatcat 120
Db 61 CTGCTCCCATCGCCATCCCCTATGCAGAGGGACAAAAGAAATACAAATTCAT 120
QY 121 gaattcaaaaaatcagcaaaagactaccctaatacaaaaatagatccagcactgaagataaaa 180
Db 121 GAATTCAAAAATCAGCAAAAGACTACCCCTAATCAAAATAGATCCAGCAGCTGAAGATAAAA 180
QY 181 accaaaaaagtgaatactgcagaccacaaatgtgcttaataatagatgtactaggaataaaagactt 240
Db 181 ACCAAAAAAGTGAATACTGCAGACCAATGTGCTAATAATAGATGTACTAGGAATAAAGGACTT 240
QY 241 ccattcacttgcaaggcttttgttttgataaaagcaagaaaaaacaatgcctctggttcccc 300
Db 241 CCATTCACTTGCAAGGCTTTTGTTTTGATAAAGCAAGAAAAACAATGCCCTCTGTTCCCC 300
QY 301 ttcaatagcatgtcaagtggagtgaaaaaagaatttggcccatgaatttgacctctatgaa 360
Db 301 TTCAATAGCATGTCAAGTGGAGTGAAAAAAGAATTTTGGCCCATGAATTTGACCTCTATGAA 360
QY 361 acaaaagactacattagaaaactgcatacttggttaaggacgcagctacaaggggaacagta 420
Db 361 AACAAAGACTACATTAGAAAACCTGCATCATTTGGTAAAGGACGCAGCTACAAGGGAAACAGTA 420
QY 421 tctatcactaagagtggcatcaaatgtcagccctggagtcccatgataccacacgaacac 480
Db 421 TCTATCACTAAGAGTGGCATCAAATGTGAGCCCTGGAGTTCCATGATACCAACACGAAACAC 480
QY 481 -----agctatcgggggtaaaagacctacaggaataactactgtcgaaatcct 525

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 2187)
AUTHORS Kitamura,N., Naka,D., Matsui,R., Yoshiyama,Y., Ishii,T. and Takahashi,K.
TITLE METHOD FOR PRODUCING HUMAN HEPATIC PARENCHYMA CELL GROWTH FACTOR AND TRANSFORMANT PRODUCING SAME FACTOR
JOURNAL Patent: JP 1991285693-A 1 16-DEC-1991;
MITSUBISHI KASEI CORP
COMMENT OS Homo sapiens (human)
PN JP 1991285693-A/1
PD 16-DEC-1991
PF 03-APR-1990 JP 1990088592
PI KITAMURA NAOMI, NAKA DAICHI, MATSUI RIE, YOSHIYAMA YOSHIKO, PI ISHII TAKEHISA, TAKAHASHI KAZUNOBU
PC C12P21/02,C12N5/10//C12N15/12.(C12P21/02,C12R1:91); CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: cell_type=serum;
FH Key Location/Qualifiers
FH
FT sig_peptide 1..87
FT sig_peptide 1..93
FT CDS 1..2187 /product='human hepatocyte growth factor'.
FEATURES
source Location/Qualifiers
1..2187
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 674 a 455 c 501 g 557 t
ORIGIN

Query Match 48.3%; Score 832.8; DB 10; Length 2187;
Best Local Similarity 98.1%; Pred. No. 9.7e-212;
Matches 859; Conservative 0; Mismatches 2; Indels 15; Gaps 1;
QY 1 atgtgggtgaccaaactcctgccagccctgctgtgcagcatgtcctcctgcctcctc 60
Db 1 ATGTGGGTGACCAAACTCCTGCCAGCCCTGCTGTGTCGAGCATGTCTCTCCTGCATCTCTC 60
QY 61 ctgctccccatcgccatccccctatgcagaggggacaaaaggaaagaataataatcat 120
Db 61 CTGCTCCCCATCGCCCATCCCTCTATGCAGAGGGACAAAAGGAAAGAATAACAAATTCAT 120
QY 121 gaattcaaaaaatcagcaaaagactaccctaatacaaaatagatccagcactgaagataaaa 180
Db 121 GAATTCAAAAAATCAGAAAGACTACCCTAATCAAAATAGATCCAGCACTGAAGATAAAA 180
QY 181 accaaaaaagtgaatactgcagaccaccaatgtgctaataagatgtactaggaaataaaggactt 240
Db 181 ACCAAAAAGTGAATACTGCAGACCAATGTGCTAATAGATGTACTAGGAATAAAGGACTT 240
QY 241 ccattcacttgcaaggcttttgtttttgataaaagcaagaaaaacaatgcctctggttcccc 300
Db 241 CCATTCACTTGCAAGGCTTTTGTTTTGTATAAAGCAAGAAAAACAATGCCTCTGGTTCCCC 300
QY 301 ttcaatagcatgtcaagtgagtggaagtgaaaaaagaatttggccatgaatttgacctctatgaa 360
Db 301 TTCAATAGCATGTCAAGTGGAGTGAAAAAAGAATTGGCCCATGAATTTGACCTCTATGAA 360
QY 361 aacaaagactacattagaaaactgcatcatttgtaaggagcgcagctacaaagggaaacagta 420
Db 361 AACAAAGACTACATTAGAAACTGCATCATTTGGTAAGGACGCAGCTACAAGGGAAACAGTA 420
QY 421 tctatcactaagagtggcacataaagtgcagccctggagttcccatgataccacacgaacac 480
Db 421 TCTATCACTAAGAGTGGCATCAAAATGTGAGCCCTGGAGTTCCTATGATACCACACGAAAC 480
QY - 481 -----agctatcgggggtaaagaccctacaggaaaaactactgtcgaaatcct 525
Db 481 AGCTTTTGGCCTTCGAGCTATCGGGGTAAAGACCTACAGGAAAAAATACTGTGGAATCCT 540

QY 526 cgaggggaagaagggggaccctggtgtttcacaaagcaatccagaggtacgctacgaagtc 585
Db 541 CGAGGGGAAGAAGGGGACCCTGGTGTTCACAAGCAATCCAGAGGTACGCTACGAAGTC 600
QY 586 tgtgacattcctcagtggttcagaagtgaattgaatgcacatgcacatggggagagttatcga 645
Db 601 TGTGACATTCTCAGTGTTCAGAAGTTGAATGCATGACCTGCAATGGGGAGAGTTATCGA 660
QY 646 ggtctcatgggatcatcacagaatcaggccaagatttgtcagcgctgggatcatcagacacca 705
Db 661 GGTCTCATGGATCATACAGAATCAGGCAAGATTGTGTCAGCGCTGGGATCATCAGACCCA 720
QY 706 caccggcacaaaattcttgcctgaaagatatccccgacaaagggctttgatgataattatgc 765
Db 721 CACCGGCACAAATCTTGCCTGAAAGATATCCCGACAAGGGCTTTGATGATAATATTATGC 780
QY 766 cgcaatcccgatggccagccgagggccatggcgctgtactcttgacccctcacacccgctgg 825
Db 781 CGCAATCCCGATGGCCAGCCGAGGCCCATGGTGCTATACTCTTGACCCCTCACACCCGCTGG 840
QY 826 gagtactgtgcaattaaaaacatgcgctgcgctgacaaaagct 861
Db 841 GAGTACTGTGCAATTAAAAACATGCGCTGACAATACT 876

Search completed: September 18, 2001, 17:44:38
Job time: 4975 sec